

Result No.	Score	Query %		DB	ID	Description
		Match	Length			
1	1770	98.8	339	1	W95578	C albicans Cdk act
2	276.5	15.4	349	1	W44767	MAP kinase. Schizo
3	261	14.6	544	1	W85028	CDK2-green floures
4	261	14.6	544	1	W85029	Green fluorescent
5	261	14.6	298	1	W95689	Homo sapiens Cdc2
6	258.5	14.4	294	1	W95690	Oryza sativa Cdc2
7	255.5	14.3	424	1	R70768	JNK2 Oncoprotein p
8	255.5	14.3	416	1	W06795	High Osmolality Gl
9	255.5	14.3	297	1	W95686	Schizosaccharomyce
10	254.5	14.2	423	1	R89409	Stress activated p
11	252.5	14.1	365	1	W55967	Human stress-activ
12	251.5	14.0	365	1	W23783	Human mitogen-acti
13	251.5	14.0	365	1	W26578	Human MAP kinase h
14	247.5	13.8	382	1	R9051	PJG4-5-CDK-BP clon
15	246.5	13.8	423	1	R89408	Stress activated p
16	246.5	13.8	365	1	W94073	Human CSBP beta po
17	246	13.7	317	1	W01559	CDK1 protein. New
18	246	13.7	317	1	W95682	Candida albicans C
19	243.5	13.6	375	1	W15512	MAP kinase #1. Wou
20	240.5	13.4	297	1	W95688	Homo sapiens Cdc2
21	238.5	13.3	297	1	W95687	Rattus norvegicus
22	236.5	13.2	384	1	R70767	JNK1 Oncoprotein p
23	236.5	13.2	633	1	W85009	Jnk1-green floures
24	236.5	13.2	630	1	W85018	Jnk1-green floures
25	235.5	13.1	323	1	W95685	Emericella nidulan
26	235	13.1	360	1	R71677	Human CSBP2. New c
27	235	13.1	360	1	W49825	Cytokine suppress
28	235	13.1	360	1	W49066	Human monocyte CSA
29	235	13.1	607	1	W85010	p38-green flouresc
30	235	13.1	606	1	W85019	p38-green flouresc
31	234.5	13.1	411	1	R89412	Stress activated p
32	234	13.1	426	1	R89410	Stress activated p
33	234	13.1	384	1	R89411	Stress activated p
34	234	13.1	393	1	W61252	Salicylic acid ind
35	233	13.0	364	1	W40061	Human MAP kinase p
36	233	12.9	364	1	W65453	Drug binding prote
37	233	12.8	360	1	R71676	Human CSBP1. New c
38	230	12.8	360	1	W49824	Cytokine suppress
39	230	12.8	360	1	W49065	Human monocyte CSA
40	227	12.7	423	1	W15513	MAP kinase #2. Wou
41	223.5	12.5	298	1	W95683	Saccharomyces cere
42	221.5	12.4	324	1	W95684	Ajellomyces capsul
43	218	12.2	393	1	R60700	Human ERK-5. Extra

RESULT	2
W44767	
ID	W44767 standard; Protein; 349 AA.
AC	W44767;

DT 16-JUN-1998 (first entry)

DE MAP kinase.

KW Mitogen activated protein; microtubule associated protein; MAP kinase;

OS cancer.

KW Schizosaccharomyces pombe.

PN J09313189-A.

PD 09-DEC-1997.

PF 29-MAY-1996; 157503.

PR 29-MAY-1996; JP-157503.

PA (KAGAKU) KAGAKU GIJUTSU SHINKO JIGYODAN.

DR WPI; 98-080080/08.

DR N-PSDB; V05891, V05892.

PT Schizosaccharomyces pombe L972 MAP kinase gene - used for the study

PT of cancer

PS Claim 1; Page 6; 12pp; Japanese.

CC This sequence is the MAP kinase of the invention. The gene

CC contributes to the elucidation of a response mechanism for the

CC environmental factors acting on cells causing cancers.

SQ Sequence 349 AA;

Query Match 15.4%; Score 276.5; DB 1; Length 349;

Best Local Similarity 26.9%; Pred. No. 5.5e-17;

Matches 93; Conservative 60; Mismatches 114; Indels 79; Gaps 15;

QY 2 KLSDYVIDKELIYNSALSDIYTAIDKFNFLPVCLKIVDEDFSLP--PHSIHREIFILKTL 59

DB 15 EITRSDLDQIGMGAGLVCSAKDLTGNAVAVKIMKPFSTPVLAFTYRELKLLKHL 74

QY 60 KPHNIIIEYFNDLKI--YDDVILVTKLYRDLSQLIEITKYCKRTTRFYINGNLVSNQ 117

DB 75 R-HENIIS-LSDIFISPFEDIYFVTELLGDLHRL--TSR----- 112

QY 118 YTLANEIEEKDKLWLSMSSGLEFTHSQIHRDKPSNIFARDITQPIIGDFICY 177

DB 112 -----PLETOFIQYFLYQLRGLKLVHSAQVHRDLKPSNILEN-----CDLKIC- 159

QY 178 DLKLPPKDEPPMAKYIDVSTGIYKAPELIIGITNYEYEDISGLIILGLYSENFSQVL 237

DB 159 DFLGLARIQDPMYGY--VSTRYRAPEIMLTWQYKYNVEVDINSAGCIF-----AEM 207

QY 238 VKDDKELTNSVDLYLLNQIENFGTNPMTDFEELFCDEY-----NNENLHFK 288

DB 208 IEGRPLPPGRDHVNOFSIETL---LGTPEMEVI--ETICKNTLRFVQSLPQKEKVPF- 262

QY 289 KFNLOKYPKRDWDILLPRCNDLDMKEIFTKWIRYDRSKRTITSKEIL 334

DB 262 ---AEKFNADPDPAI-----DLLEKMLVDFPKRISAADAL 294

RESULT 3

W85028

ID W85028 standard; Protein; 544 AA.

AC W85028;

DT 08-FEB-1999 (first entry)

DE CDK2-green fluorescent protein union product.

KW Human; CDK2 gene; fusion protein; green fluorescent protein; GFP;

KW intracellular signalling; chimera.

OS Chimeric - Aequorea victoria.

OS Chimeric - Homo sapiens.

PN W09845704-A2.

PD 15-OCT-1998.

PF 07-APR-1998; DK0145.

PR 07-APR-1997; DK-000392.

PA (NOVO) NOVO-NORDISK AS.

PI Kasper A, Petersen Bjorn S, Scudder K, Thastrup O,

PI Tullin S;

DR WPI; 98-594491/50.

DR N-PSDB; V71073.

PT Determining effect on signalling pathways in live cells from

PT redistribution of luminophores - specifically fusions of green

PT fluorescent protein with a signalling component, and new apparatus,

PT particularly for identifying toxins and potential therapeutic agents

Query Match 14.6%; Score 261; DB 1; Length 544;

Best Local Similarity 26.2%; Pred. No. 2.4e-15;

Matches 83; Conservative 59; Mismatches 119; Indels 56; Gaps 11;

QY 21 IYTAIDKFNFLPVCLKIV--DEDFSLPHPSIHRIFILKTLKPHNPNIIEYFNDLKIYDDV 78

DB 18 VYKARNKLTGEVVALKIRLDTEGVPSSTAIREISLLKELN-HPNIVKLLDVIHTENKL 76

QY 79 ILVTKLYRDLSQLIEITKYCKRTTRFYINGNLVSNQYTLANEIEEKDKLWLSMSS 138

DB 77 YLVFPEFRLQDLKKFMDASA-----LTGIPLP-----IKSYLFQQLQ 113

QY 139 GLEFIHSQIHRDKPSNIFARDITQPIIGDFICYDLKLPPKDEPPMAKYI-DVST 197

DB 114 GLAFCHSRHLVHRDLKPNQLLINTEGAIK--LADFLARAFV-----PVRTYTHEVVT 165

QY 198 GIYKAPELIIGITNYEYEDISGLIILGLYSENFSQVLVKDDKELTNSVDLYLLN 257

DB 166 LWYRAPEILLGLSKYISTAVDWSLGCIF-----AEMVTRRALFPDGDSEIDQLF--- 214

QY 258 QIFENFGTNPMTDFEELFCDEYNENLHFKFNLOKYPKRDWDIILPRCNDLDMKEIFT 317

DB 214 RIIFTLGTPEVWVPGVTSMPLY-----KPSFPRKARQDFSKVVPPLDED-GRSLLS 264

QY 318 KMIRYDRSKRTITSKEIL 334

DB 265 QMLHYDPNKRISAKAL 281

RESULT 4

W85029

ID W85029 standard; Protein; 544 AA.

AC W85029;

DT 08-FEB-1999 (first entry)

DE Green fluorescent protein-CDK2 fusion product.

KW Human; CDK2 gene; fusion protein; green fluorescent protein; GFP;

KW intracellular signalling; chimera.

OS Chimeric - Aequorea victoria.

OS Chimeric - Homo sapiens.

PN W09845704-A2.

PD 15-OCT-1998.

PF 07-APR-1998; DK0145.

PR 07-APR-1997; DK-000392.

PA (NOVO) NOVO-NORDISK AS.

PI Kasper A, Petersen Bjorn S, Scudder K, Thastrup O,

PI Tullin S;

DR WPI; 98-594491/50.

DR N-PSDB; V71074.

PT Determining effect on signalling pathways in live cells from

PT redistribution of luminophores - specifically fusions of green

PT fluorescent protein with a signalling component, and new apparatus,

PT particularly for identifying toxins and potential therapeutic agents

PS Example 12; Pages 212-213; 326pp; English.

CC The present sequence represents a green fluorescent protein (GFP)-human

CC CDK2 fusion protein. The fusion protein is used in an assay to

CC exemplify the invention. The specification describes how quantitative

CC information about the influence of a molecule on a cellular response is

Best Local Similarity 25.3%; Pred. No. 1.7e-15;
Matches 85; Conservative 60; Mismatches 132; Indels 59; Gaps 12;

```
QY 5 DYVIDKELIYNSAISIITATDKNNLPVCLK---IYDEDFSLPPHSIHREIFILKTKP 61
Db 2 EQYEKEKIGEGTYGVTRADKVTNETIALUKIRLQSEDEGFTAI-REISLKEH- 60
QY 62 HPMIEFYNDLKIYDVLVTKLYRDLSQLIE-ITKYCKRTTRFYINGNLVSNQYTL 120
Db 60 HGNIVRLHDVHSEKRIYLVFEYLDLKKFMDSCPEAKNPDL----- 104
QY 121 ANEIEBKDKLWKLKSSGSLFIHSGIITHRDIKPSNIFFARDDITQPI-IGDFICIDL 179
Db 104 -----IKSYLIQILGVAYCHSRVHLRDLAPQLLI--DRTNALKADFGARAF 153
QY 180 KLPPKDEPPMAKYIDVSTGIYKAPELILGTINYEYEDISLGIILTYLSENFOSVLVK 239
Db 154 GIPVRFTH-----EWTILWYRAPEILIGSRQYSTPVDMSVGCIF-----AEMVN 199
QY 240 DKELTNDHVSDDLKLLQIENFGTNPITDFEELFCDEYNNENLHFKFNLOKYPKO 299
Db 200 QKLPFGSEIDELF---KIFVLGTNPEQSGVGSLLPYI-----KSAPPRWQAO 248
QY 300 WDIILPRCNDLMLKEIFTKMIRYDRSKRITSKEILQ 335
Db 249 LATIIVPTL-DPAGLLOLLSKMLRYEPNKRITARQALE 283
```

RESULT 7

```
ID R70768 standard; Protein; 424 AA.
AC R70768;
DT 13-OCT-1995 (first entry)
DE JNK2 Oncoprotein polypeptide kinase.
KW Oncoprotein; polypeptide kinase; JNK; c-Jun; AP-1; leukaemia;
KW prostate cancer; lung cancer; colon cancer.
OS Homo sapiens.
PN W0950323-A.
PD 02-FEB-1995.
PR 18-JUL-1994; U08119.
PR 19-JUL-1993; US-094533.
PR 25-MAR-1994; US-020602.
PA (REGC) UNIV CALIFORNIA.
PA (UYMA-) UNIV MASSACHUSETTS MEDICAL SCHOOL.
PI Davis R, Derlward B, Hibbi M, Karin M, Lin A;
PI WPI; 95-075186/10.
DR N-PSDB; Q85309.
DR Oncoprotein protein kinase, JNK - useful for regulating c-Jun
PT activity and gene expression in cells
PS Clajm 55; Figure 29; 143pp; English.
CC activation domain which affect gene expression from AP-1 sites. The
CC 2 protein kinases allow the detection of the level of specific
CC kinase activity associated with activation of c-Jun and AP-1. Cell
CC proliferative disorders associated with JNK can be treated by
CC administering a therapeutically effective amount of a reagent which
CC modulates JNK activity. Disorders such as leukaemia, cancers of
CC prostate, lung (non-small cell) and colon.
SQ Sequence 424 AA;
```

Query Match 14.3%; Score 255.5; DB 1; Length 424;

Best Local Similarity 27.1%; Pred. No. 5.3e-15;
Matches 92; Conservative 49; Mismatches 112; Indels 87; Gaps 16;

```
QY 21 IYTAIDKFNLPVCLKIVDEDFSLPPHS--THREIFILKTKPNNIIEYFN-----DL 72
Db 40 VCAAPTVLGISVAVKLSRPFQNTAKRAYRELVLKCVN-HKNILSLNWFIPQKTL 98
QY 73 KIYDDVLVTKLYRDLSQLIEITKYCKRTTRFYINGNLVSNQYTLANEIEKDIKLM 132
Db 99 EEFQDYLLYMLMDANLCQVIHM-----ELDHRSYSL 131
```

```
QY 133 LKMSGSLFEHISQIITHRDIKPSNIFFARDDITQPIIGDFDI-----CYDLKLPPKDEPP 188
Db 132 LYOMLCGIKHLSAGIITHRDLKPSNI-VKSDCTLKIL-DFGLARTACTNFWMTFY---- 186
QY 189 MAKYIDVSTGIYKAPELILGTINYEYEDISLGIILTYLSENFOSVLVKDKKELTND 248
Db 186 -----VTRYRAPEVLJGM-GYKENVDIWSVGCIM-----GELVKGCVIFQGT 229
QY 249 HVSDDLKLLQIENFGTNPITDFEELFCDEYNNENLHFKFNLOKYPK-----KMD 301
Db 230 HIDQ---WNKVIEQLGTPS-AEPMKKLOPTVRNVE-----NRPKYPGIKFEELPDW- 279
QY 302 IILP-----RCNDLMLKEIFTKMIRYDRSKRITSKEILQ 335
Db 279 -IFPSESEBDIKITSQARDLLSKMLVIDPKRISVDEALR 317
```

RESULT 8

```
W06795
ID W06795 standard; Protein; 416 AA.
AC W06795;
DT 27-JAN-1997 (first entry)
DE High Osmolality Glycerol 1 gene product.
KW HOG1; hog1; hog2; mutant; glycerol; production; reduce; baking;
KW beverage; yeast; fermentation.
OS Saccharomyces cerevisiae.
PN US5545556-A.
PD 13-AUG-1996.
PF 23-MAY-1991; 704943.
PR 23-MAY-1991; US-704943.
PR 18-MAR-1993; US-032382.
PA (UYRI-) UNIV RICE WILLIAM MARSH.
PI Gustin MC;
PI WPI; 96-383673/38.
DR N-PSDB; T44328.
DR Yeasts contg. mutated HOG gene or multiple copies of HOG gene - have
PT reduced or increased glycerol production
PS Example 5; Fig 5; 28pp; English.
CC Low glycerol fermentation products can be produced using mutant strains
CC of yeast, e.g. Saccharomyces sp. or Torulaspora sp., that have a reduced
CC glycerol prodn. relative to the unmutated parental strain. High
CC Osmolality Glycerol genes have been identified, specifically HOG1, HOG2
CC and HOG4. Yeast strains contg. mutant HOG genes have a reduction in
CC glycerol biosynthesis and arg. useful in prodn. of novel beverages.
CC baked goods or glycerol. HOG1 and HOG2 are defined as yeast genomic DNA
CC segments capable of complementing the reduced growth in high osmolality
CC media (defined as the Osm phenotype) and glycerol biosynthesis defect
CC of the hog1 and hog2 mutants, respectively. Southern blot hybridisation
CC of HOG1 to electrophoretically separated S. cerevisiae chromosomes showed
CC hybridisation of HOG1 to chromosome XII. The present sequence is that of
CC the HOG1 gene product of 416 amino acids with a mol. wt. of 47 kDa.
SQ Sequence 416 AA;
```

Query Match 14.3%; Score 255.5; DB 1; Length 416;

Best Local Similarity 23.1%; Pred. No. 5.1e-15;
Matches 84; Conservative 63; Mismatches 103; Indels 113; Gaps 13;

```
QY 2 KLSDDYIDKELIYNSAISDIYTAIDKFNLPVCLKIVDEDFS--LPPHSIHREIFILKTL 59
Db 18 EITNRYNDLPVKGAGLVCASATDTLTSQPAVKIKMPFSTAVLAKRYRELKLLKHL 77
QY 60 KPHPNII---EYFNDLKYDVLVTKLYRDLSQLIEITKYCKRTTRFYINGNLVSN 116
Db 78 R-HENLICLQDIF--LSPDIEYFVTELOGTDLRLIQ----- 113
QY 117 QYTLANEIEKDIKLTWKLKSSGSLFIHSGIITHRDIKPSNIFFARDDITQPIIGDFDIC 176
Db 113 ----TRPLEKQFVQYFLYQILRGLKYVHSAGVHRDLKPSNILEN-----CDLKIC 161
QY 177 YDLKLPPKDEPPMAKYIDVSTGIYKAPELILGTINYEYEDISLGIILTYLSENFOSV 236
Db 162 -DFGLARIQDPQMTGY--VSTRYIRAPEILMTWQKYDVEVDIWSAGCIF-----AE 209
```


GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 6, 1999, 08:13:57 ; Search time 16.13 Seconds
(without alignments)
1293.447 Million cell updates/sec

Title: US-09-072-994-14

Perfect score: 1791

Sequence: 1 MKLSDYIDKELIYNSAISD.....INYSKRITSKEILQLMLD 339

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

Database: SPREMBL_10.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	271.5	15.2	311	5	O15851	O15851 leishmania
2	271.5	15.2	311	5	O96526	O96526 leishmania
3	268	15.0	288	5	O96821	O96821 plasmodium
4	264.5	14.8	407	3	O59854	O59854 zygocacchar
5	264.5	14.8	294	10	O40484	O40484 nicotiana t
6	263.5	14.7	311	5	O94796	O94796 trypanosoma
7	263.5	14.7	294	10	O65838	O65838 lycopersico
8	262	14.6	391	11	O61177	O61177 mus musculus
9	261	14.6	298	5	O26671	O26671 theileria a
10	260	14.5	307	5	O27168	O27168 paramacium
11	259.5	14.5	382	13	P79996	P79996 gallus gall
12	259	14.5	391	4	O14013	O14013 homo sapien
13	258.5	14.4	366	5	O62618	O62618 drosophila
14	257.5	14.4	380	3	O93982	O93982 zygocacchar
15	257.5	14.4	294	10	O40482	O40482 nicotiana t
16	257.5	14.4	302	10	O38772	O38772 antirrhinum
17	257	14.3	298	5	O27032	O27032 theileria p
18	257	14.3	508	5	O25917	O25917 plasmodium
19	256.5	14.3	765	5	O25758	O25758 plasmodium
20	256.5	14.3	826	5	O94656	O94656 plasmodium
21	255	14.2	307	10	O04402	O04402 petunia hyb
22	255	14.2	369	10	O43379	O43379 avena sativ
23	255	14.2	302	10	O40734	O40734 oryza sativ
24	254.5	14.2	294	10	O22292	O22292 allium cepa
25	253.5	14.2	294	10	P93101	P93101 chenopodium
26	253.5	14.2	371	10	O04694	O04694 petroselinu
27	253	14.1	288	5	O96820	O96820 plasmodium
28	252.5	14.1	365	4	O15264	O15264 homo sapien
29	252.5	14.1	294	10	O82135	O82135 pisum sativ

30	252	14.1	346	11	O60545	O60545 mesocricetu
31	251.5	14.0	366	5	O46216	O46216 drosophila
32	251.5	14.0	294	10	O65839	O65839 lycopersico
33	251	14.0	293	10	O40483	O40483 nicotiana t
34	250	14.0	350	11	O54833	O54833 mus musculu
35	249	13.9	369	10	O81599	O81599 triticum ae
36	248	13.8	300	5	O44000	O44000 toxoplasma
37	248	13.8	377	5	O17446	O17446 caenorhabdi
38	248	13.8	346	11	O55077	O55077 cricetus
39	248	13.8	427	13	O42099	O42099 cyprinus ca
40	247.5	13.8	420	3	O59853	O59853 zygocacchar
41	247.5	13.8	365	4	O14739	O14739 homo sapien
42	247.5	13.8	294	10	O41639	O41639 vigna aconit
43	247	13.8	346	11	O55105	O55105 mus musculu
44	246.5	13.8	365	4	O15124	O15124 homo sapien
45	246.5	13.8	294	10	O40789	O40789 petroselinu

ALIGNMENTS

RESULT 1

O15851
ID O15851 PRELIMINARY; PRT; 311 AA.
AC O15851;
DT 01-JAN-1998 (TEMBLrel. 05, Created)
DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)
DT 01-MAY-1999 (TEMBLrel. 10, Last annotation update)
DE CDC2-RELATED KINASE 3.
GN CRK3.
OS Leishmania mexicana.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 98221140.
RA GRANT K.M., HASSAN P., ANDERSON J.S., MOTTRAM J.C.;
RT "The crk3 gene of Leishmania mexicana encodes a stage-regulated
cdc2-related histone H1 kinase that associates with p12";
RL J. Biol. Chem. 273:10153-10159(1998).
DR EMBL; AJ001275; CAA04648.1; -.
DR PFAM; PF00069; pkinase; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
SQ SEQUENCE 311 AA; 35584 MW; 35G4E98D CRC32;

Query Match 15.2%; Score 271.5; DB 5; Length 311;
Best Local Similarity 26.9%; Pred. No. 1.9e-11;
Matches 94; Conservative 53; Mismatches 117; Indels 85; Gaps 14;

QY 5 DYIDKELIYNSAISDIYTAIDKFNLPVCLKTV--DEDFSLPPHSIHREIFLTKLKP 62
Db 21 DRYNLDVLGEGYGVVYRAVDKITQYVALKKVRLDRTEEGIPQALREVSLQEF-DH 79
QY 63 PNIEIF-----NDLKYYDDVILVKLYRVDLSOLIEITKYCKRTTFYINGNLVSN 118
Db 80 PNIVNLDVICSDGKLY----LVFEYVEADLKKAIEKQE-----GGYSG----- 120
QY 119 TLANEIEEKDKLWLKSMSSGLEFIHSGIHRDKPSNIFFARDDITQPIGDFDICVD 178
Db 120 -----MDLKRLLIQLDGLGYFCHRRHRIHRDLKPANILTSGNVLK--LADFGLARA 169
QY 179 LKLPKPDEPPMAKYI-DVSTGIYKAPLEILGITNYEYEDIDWSLGIILTGLYSENFQSVL 237
Db 170 FQV-----PMHTYTHEVYTLWRAPEILLGKHPAPVDMKMSVGCIFAELAR---RKVL 220
QY 238 VRDDKELTNDSHVSDLYLLNQIFENFGTP-----NLTDFEDELFCDEYNNENLH 286
Db 221 FRGDSIEGQ-----LFEIFQVLGTPTDTEGSGWPGVSRPLPYRDV----- 260
QY 287 FKKFNLQKYPKRDWDIILPRCNDLMLKEIFTKMIRYDRSKRITSKEILQ 335
Db 260 -----FPKWTAKRLQVLPDELHPDAI-DLLSKMLKYPDRISAKALQ 302

```
RESULT 2
O96526 ID O96526 PRELIMINARY; PRT: 311 AA.
AC O96526;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-MAY-1999 (Tremblrel. 10, Last annotation update)
DE CDC2-RELATED KINASE.
GN CRK3.
OS Leishmania major.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FRIDLIN VI.
RA WANG Y., DIMITROV K., GARRITY L.K., SAZER S., BEVERLEY S.M.;
RT "Stage-specific activity of Leishmania major CRK3 kinase and
RT functional rescue of a Schizosaccharomyces pombe cdc2 mutant.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF073381; AAD08994.1; -
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Kinase.
SQ SEQUENCE 311 AA; 35644 MW; C7716C68 CRC32;

Query Match 15.2%; Score 271.5; DB 5; Length 311;
Best Local Similarity 26.9%; Pred. No. 1.9e-11;
Matches 94; Conservative 53; Mismatches 117; Indels 85; Gaps 14;

QY 5 DYIDKELIYNSAIDYTAIDKFNLPVCLKIV--DDFSLPHSHIHFILKTLKPH 62
DB 21 DRYNLDVLGEGTYGVYVAVDKITGTQVALKKVLDRTEGIPOTALREVSILQEF-DH 79
QY 63 PNIEYF---NDLKYYDVLVTLKLYDLSOLSEIFEIKYCKRTRETYINGNLVSNQY 118
DB 80 PNIVNLLDVCSDGLY---LVFEYVADLKAIEKQE-----GGYSG----- 120
QY 119 TLANEETEEDKILWLKSMSSGLEFTHSGIHRDKIPSNIEFFARDDITQPIIGDFICYD 178
DB 120 -----MDLRLIYQLLDGLYFCHRRHRIHRDLKPNALLTSGNVLK--LADFGIARA 169
QY 179 LKLPPKDEPPMAKYI-DVSTGIYKAPELILGITYEYEDISGLITGLYSENFOSVL 237
DB 170 FQV-----PMITYTHEVVTLYWRAPEILLGKHVTPAIDVMSVGCIFAEEL---TRKVL 220
QY 238 VKDDKELTNDSHVSLYLNLQIFENFGTP-----NLTFDEFLCFDEYNNENLH 286
DB 221 FRGDSIEGO-----LFEIFQVLGTPDTDEGSPGVSRLPDYDV----- 260
QY 287 FKKFNLOKYPKRDWILLPRCNDLMLKEIFTKMIYDRSKRITSKEILQ 335
DB 260 -----FPKWATKRLGOVLPELHPDAI-DLLSKMLKIDPRERISAKEALQ 302

RESULT 3
O96821 ID O96821 PRELIMINARY; PRT: 288 AA.
AC O96821;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-MAY-1999 (Tremblrel. 10, Last annotation update)
DE CDC2-RELATED KINASE 2.
OS Plasmodium knowlesi.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LINE H;
RX MEDLINE; 96065755.
RA VINKENOOG R., VELDHUISSEN B., SPERANCA M.A., DEL PORTILLO H.A.,
RA JANSE C.J., WATERS A.P.;
RT "Comparison of introns in a cdc2-homologous gene in a number of
RT Plasmodium species.";
RL Mol. Biochem. Parasitol. 71:233-241(1995).
```

```
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=LINE H;
RA VINKENOOG R., SPERANCA M.A., RAMESAR J., THOMAS A.W.,
RA DEL PORTILLO H.A., JANSE C.J., WATERS A.P.;
RT "Characterisation of the cdc2-related kinase 2 gene from Plasmodium
RT knowlesi and P. berghei.";
RL Mol. Biochem. Parasitol. 95:229-240(1998).
DR EMBL; AJ224155; CAAL1852.1; -
KW Kinase.
SQ SEQUENCE 288 AA; 32946 MW; 2E54A477 CRC32;

Query Match 15.0%; Score 268; DB 5; Length 288;
Best Local Similarity 27.5%; Pred. No. 2.9e-11;
Matches 84; Conservative 55; Mismatches 103; Indels 64; Gaps 11;

QY 35 LKIVDEDFSLPHSHIHFILKTLKPHNIEVFNDLKIYDDVILVTLKLYRYDL3QLIE 94
DB 34 IRLKEDEGIPSTAI-REISILKELK-HSNIVKLYDVHTKKLLLVFEHLDDQDLAKLD 91
QY 95 ITKYCKRTREIYINGNLVSNQYTLANEETEEDKILWLKSMSSGLEFTHSGIHRDIK 154
DB 92 V-----CDGGL-----ESVTAKSFLQLLSGIAYCHEHRLVHDLK 127
QY 155 PSNIEFFARDDITQPIIGDFICYDILKLPKDEPPMAKYI-DVSTGIYKAPELILGITYE 213
DB 128 PQNLLINREGELK--IADFGIARAFGI-----PVRKYTHEVVTLYWRAPDILMGSKYS 179
QY 214 YEIDWISGLIITGLYSENFOSVLVKDDKELTND----SHVSDLYLLNLQIFENFGTPNLT 269
DB 180 TPIDWISVGCIFA-----EMVNGRPLFCVSETDQLMIRIFRLIGTPNSA 223
QY 270 DFEDELFCDEYNNENLHFKFNKLYKQYPRKDWIILPRCNDLMLKEIFTKMIYDRSKRIT 329
DB 224 NWPSPVTELPKYPDFI-----VYEPLPWEITFL-KGLDDTGIDLLSKMLRLDPNQRIT 274
QY 330 SKEILQ 335
DB 275 AKEALQ 280

RESULT 4
O59854 ID O59854 PRELIMINARY; PRT: 407 AA.
AC O59854;
DT 01-AUG-1998 (Tremblrel. 07, Created)
DT 01-AUG-1998 (Tremblrel. 07, Last sequence update)
DT 01-MAY-1999 (Tremblrel. 10, Last annotation update)
DE ZHOG1P.
GN ZHOG1.
OS Zygosaccharomyces rouxii (Candida mogii).
OC Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;
OC Saccharomycetaceae; Zygosaccharomyces.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC42981;
RA WATANABE Y.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB012146; BAA25200.1; -
DR PFAM; PF00069; pkinase; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
SQ SEQUENCE 407 AA; 46428 MW; 56BDD207 CRC32;

Query Match 14.8%; Score 264.5; DB 3; Length 407;
Best Local Similarity 23.7%; Pred. No. 7.8e-11;
Matches 86; Conservative 61; Mismatches 103; Indels 113; Gaps 13;

QY 2 KLSDYIDKELIYNSAIDYTAIDKFNLPVCLKIVDEDFS--LPPHSIHRHIFIKTL 59
DB 18 EITNRYTDLNPVGAGFLVCSATDILAGQPVAKIMKPFSTAVLAKRTYRELKLLKHL 77
```

```

QY 60 KPENII---EYNDLKIYDDVILVTKLYRDLSQLIEITKYCKRTRTFYINGNLVSN 116
Db 78 R-HENLICLODIF--LSPLEDIVFVTELOQDLHRLQ----- 113
QY 117 QYTLANEIEKDKLWLKSSGLFIHSHOGIIRHDIKPSNIFFARDDITQPIIGDFDIC 176
Db 113 ---TRPLEKQFOYFELYQIRLGLKVHSGVIRHDLKPSNLINEN-----CDLKIC 161
QY 177 YDLKLPKPEPPMAKYIDVSTGIYKAPELITGINTNYEYDIWSLGIILTGLYSENFOSV 236
Db 162 -DFGLARIQDPQMGY--VSTRYIRAPEIMLTWQYDVEVDWSAGCIF-----SE 209
QY 237 LVKDDKELTNDSHVSDYLILNQIFENFGTNPNTDFEELFCDEYNNENLHFKFNLYKP 296
Db 210 MIEGKLPFPKGVHVOFSITDL---LGSP-----P 237
QY 297 RKWDIILPRCNDLAK-----EIFTKMIRYDRSKRITSK 331
Db 238 R---DVIITICSEDTLKFVTSPLRHPDVPFQERKAVEPDAVDLLGRMLVDPKKRITAA 294
QY 332 EIL 334
Db 295 DAL 297

RESULT 5
Q040484 PRELIMINARY; PRT; 294 AA.
AC Q040484;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
DE CDC2 HOMOLOG.
OC Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
OC core eudicots; Asteridae; euasterids I; Solanales; Solanaceae;
OC Nicotiana.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SAMSON NN; TISSUE=APICAL MERISTEM;
RX MEDLINE; 96245449.
RA SETIADY Y.Y., SEKINE M., HARIUCHI N., KOUCHI H., SHINYO A.;
RT "Molecular cloning and characterization of a cDNA clone that encodes
RT a Cdc2 homolog from Nicotiana tabacum.";
RL Plant Cell Physiol. 37:369-376(1996).
DR EMBL; D50738; BAA09369.1; -.
DR MENDEL; 13988; NICta:2321; mn13988.
DR PFAM; PF00069; pkinase; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
SQ SEQUENCE 294 AA; 33917 MW; 3EAC84A9 CRC32;

Query Match 14.8%; Score 264.5; DB 10; Length 294;
Best Local Similarity 25.9%; Pred. No. 5.2e-11;
Matches 89; Conservative 57; Mismatches 123; Indels 75; Gaps 13;

QY 5 DYIYDKELIYNSAISDIYTAIDFENLPLVCLK---IVDEDFSLPSPHSIHRFIFILTKLP 61
Db 2 DQYKVEKIGEGYGVVYKARDRTNETALKKIRLEQDEGVPTAI-REISLLKEMQ- 60
QY 62 HPNIIEYNDLKIYDDVILVTKLYRDLSQLIEITKYCKRTRTFYINGNLVSNQYTLA 121
Db 60 HANIV-----RLQDVVHSEKRLYLVEYLDLKKHMDSSPEF----- 98
QY 122 NEIEKDKIWLKSSGLFIIHSHOGIIRHDIKPSNIFFARDDITQPI-IGDFDICYDLK 180
Db 98 -SEDPRLVKFLYQIRLGLKVHSGVIRHDLKPSNLI--DRRTNAKLADFGLARAFG 154
QY 181 LPPKDEPPMAKYIDVSTGIYKAPELITGINTNYEYDIWSLGIILTGLYSENFQSVLVKD 240
Db 155 IPVRTFTH----EVLTLWYRAPEILLGSRHSTPTVDVWSVGCIF-----AEMVTQ 200

```

```

QY 241 DKELTNDSHVSDYLILNQIFENFGTNP-----LTDFEELFCDEYNNENLHFKFN 291
Db 201 RPLFPDSEIDELF--KIFRVMGTPNEDTWPGVTTLPDF-----KSA 240
QY 292 LQYPRKDDIILPRCNDLAKIEFTKMIRYDRSKRITSKEILQ 335
Db 241 FPKWPSKDLATIVPNL-DGAGLDLDDKMLRLDPSKRITARNAL 283

RESULT 6
Q04796 PRELIMINARY; PRT; 311 AA.
AC Q04796;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
DE CDC2-RELATED PROTEIN KINASE 1.
GN TCRK1.
OS Trypanosoma cruzi.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TUL2;
RX MEDLINE; 98226177.
RA GOMEZ E.B., KORNBLIHT A., TELLEZ-INON M.T.;
RT "Cloning of a cdc2-related protein kinase from Trypanosoma cruzi that
RT interacts with mammalian cyclins";
RL Mol. Biochem. Parasitol. 91:337-351(1998).
DR EMBL; U69958; AAC48317.1; -.
DR PFAM; PF00069; pkinase; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
SQ SEQUENCE 311 AA; 35483 MW; 8540A103 CRC32;

Query Match 14.7%; Score 263.5; DB 5; Length 311;
Best Local Similarity 25.9%; Pred. No. 6.5e-11;
Matches 87; Conservative 59; Mismatches 99; Indels 91; Gaps 16;

QY 21 IYTAIDKNNLPVCLKIV--DEDFSLPSPHSIHRFIFILTKLPHPNIEYFN---DLKI 74
Db 37 VYRAVDKITGQIVALKKVLRLDTESGIPTALREVSLQIEIH-HPNVNLLDVICTDGL 95
QY 75 YDDVILVTKLYRDLSQLIEITKYCKRTRTFYINGNLVSNQYTLANEIEKDKIWLK 134
Db 96 Y----LIFEYDYDLKKAIE-----KRGCTFT-GVT-----LKKLVY 127
QY 135 SMSGLEFIHSHOGIIRHDIKPSNIFFARDDITQPIIGDFICYDLKLPKDEPPMAKYI- 194
Db 128 QLLDGLFFCHRRHIVHSDLKPANILITSDNVLK--LADFGARTFQI-----PMHTYTH 179
QY 194 DVSTGIYKAPELITGINTNYEYDIWSLGIILTGLYSENFOSVLVKDDKELTNDSHVSD 253
Db 180 EYVTLWYRAPEILLGKHYTPAVDIWSVGCIF-----AELARGKVIKPRGDSGQL 230
QY 254 YILNQIFENFGTNP-----NLTFEELFCDEYNNENLHFKFNLOKQYPR---KD 299
Db 231 F---EIFQLITPMDNESGWPVGSLLPYRD-----VFFRWAGKP 267
QY 300 WDILLPRCNDLAKIEFTKMIRYDRSKRITSKEILQ 335
Db 268 LAQVIPQLDSEAI-DLISRLKYSAPERISAKALQ 302

RESULT 7
O63838 PRELIMINARY; PRT; 294 AA.
ID O63838;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
DE CYCLIN-DEPENDENT PROTEIN KINASE P34CDC2.
GN CDC2A-1.
OS Lycopersicon esculentum (Tomato).

```



```
Qy 65 IIEFNDKIYDDVILVTKLYRYDLSQLEIHTKYKRTTRFYIYINGNLVSNQYTLANEI 124
Db 94 IITLADIVK--DV-----SRTPALVFEHVNNITFKQ--LYQTF 128

Qy 125 EEKDKLWLMSSGLEFIHSGIHRDIKPSNIFFARDDITQPIIGDFDICYDKLPPK 184
Db 129 TDYDIRFYETIKALDYCHSGIMHROVKPHNVIMIDHEHKKRLI---DWGLAEF 191

Qy 185 DEPPMAKYIDVSTGYKAPELIGITNTVEYEDISGLIITGLY--SENFQSVLVKDDK 242
Db 182 YHPQOEYNVASRYKFKPELLVDYQMDYSLDMWSLGLASMFREKPE-----233

Qy 243 ELINDSHVSDLY--LLNQIENFGTNPNTDFEDELFCDEYNNENLHFKFN--LOKYPK 299
Db 233 -----FHGHNDYDOLVRIAKVITGLY-----YIDKINIE--LDPRFNDILGRHSKR 280

Qy 300 WDIILPCNDLDMK-----EIFTMIRYDRSKRITSKEILQ 335
Db 281 WERFVHREHQLVSPALDFLDKLLRYDHQSRLTAREAME 320

RESULT 13
ID O62618 PRELIMINARY; PRT; 366 AA.
AC O62618;
DT 01-AUG-1998 (TEMBLrel. 07, Created)
DT 01-AUG-1998 (TEMBLrel. 07, Last sequence update)
DE P38A MAP KINASE.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea;
OC Drosophilidae; Drosophila.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 98252940.
RA HAN Z.S., ENSLEN H., HU X., MENG X., WU I.-H., BARRETT T., DAVIS R.J.,
RA IP Y.T.;
RT "A conserved p38 mitogen-activated protein kinase pathway regulates
RT Drosophila immunity gene expression."
RL Mol. Cell. Biol. 18:3527-3539(1998).
DR ENBL; AFO35547; AAC39031.1; -.
DR ENBL; AFO35546; AAC39030.1; -.
DR FLYBASE; FBgn0015765; Mpk2.
DR PFAM; PF00069; pkinase; 1.
SQ SEQUENCE 366 AA; 42255 MW; F6E3EED CRC32;

Query Match 14.4%; Score 258.5; DB 5; Length 366;
Best Local Similarity 25.0%; Pred. No. 1.7e-10;
Matches 86; Conservative 61; Mismatches 130; Indels 67; Gaps 11;

Qy 2 KLSDYIDKELIYNSAISDIYTAIDKFNLPVCLKIVDEDFSLPPHS--IHREIFILKTL 59
Db 20 EIPDIYDOLQVSGGAYGVQSAVVRGTMHVAIKKLARPFQSAVHAKRTYRELRLKHM 79

Qy 60 KHPNII--EYFNDDVILVTKLYRYDLSQLEIHTKYKRTTRFYIYGIN 110
Db 80 -DHENVIGLIDIFHPANGSLENFQVYLVTHLMDADLNIRM-----124

Qy 111 GNLVSNQYTLANEIEKDKLWLMSSGLEFIHSGIHRDIKPSNIFFARDDITQPII 170
Db 124 -----QHLSDHDFVLYQILRLGLKYIHSAGVTHRDLPKPSNI-AVNEDCELRL 171

Qy 171 GDFDICYDLKPPKDEPPMAKYIDVSTGYKAPELIGITNTVEYEDISGLIITGLYS 230
Db 172 -----DFGLAPRTENETGY--VATRYRAPEIMLNWMDYDQTVLWSVGCIMAEIIT 222

Qy 231 ENFQSVLVKDDKELTNDSHVSDLYLLNQIENFGTNPNTDFEDELFCDEYNNENLHFKF 290
Db 223 RR-----TLFPQGDTHIQNLINEMLGTTPPAE-----FLKKISSARSYIQ 264

Qy 291 NLQYPRKWDIILPCNDLDMKEIFTMIRYDRSKRITSKEIL 334
```

```
Db 265 SLPPMGRGSKFNKFNAN-PLAIDLEKMLEDAEKRTAEAL 307

RESULT 14
ID O93982 PRELIMINARY; PRT; 380 AA.
AC O93982;
DT 01-MAY-1999 (TEMBLrel. 10, Created)
DT 01-MAY-1999 (TEMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TEMBLrel. 10, Last annotation update)
DE PUTATIVE MAP KINASE.
GN HOG1
OS Zygosaccharomyces rouxii (Candida mogii).
OC Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;
OC Saccharomycetaceae; Zygosaccharomyces.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CBS 732;
RA KINLOVA O., SYCHROVA H.;
RT "Cloning and sequencing of the Zygosaccharomyces rouxii HOG1 gene."
RT Submitted (Feb-1999) to the EMBL/Genbank/DBJ databases.
DR ENBL; AJ132606; CAAL0714.1; -.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Kinase.
SQ SEQUENCE 380 AA; 43298 MW; 7354C046 CRC32;

Query Match 14.4%; Score 257.5; DB 3; Length 380;
Best Local Similarity 23.4%; Pred. No. 2.1e-10;
Matches 85; Conservative 61; Mismatches 104; Indels 113; Gaps 13;

Qy 2 KLSDYIDKELIYNSAISDIYTAIDKFNLPVCLKIVDEDFSLPPHSIHREIFILKTL 59
Db 18 EITNRYTDLPVGVGAGLVCSATDTLAGQVPAIKKIMKPFSTAVLAKRTYRELKLLHL 77

Qy 60 KHPNII--EYFNDDVILVTKLYRYDLSQLEIHTKYKRTTRFYIYINGNLVSN 116
Db 78 R-HENLICQDIF--LSPLEDIYFVTEQGLDLHRLQ-----113

Qy 117 QYTLANEIEKDKLWLMSSGLEFIHSGIHRDIKPSNIFFARDDITQPIIGDFDIC 176
Db 113 ----TRPLEKQFVQVLYQILRLGLKYVHSAGVTHRDLPKPSNINEN-----CDLKIC 161

Qy 177 YDLKLPKDEPPMAKYIDVSTGYKAPELIGITNTVEYEDISGLIITGLYSENFQSV 236
Db 162 -DFGLARIQDPQMTGY--VSTRYRAPEIMLTWQKYDVEYDWSAGCIF-----SE 209

Qy 237 LVKDDKELTNDSHVSDLYLLNQIENFGTNPNTDFEDELFCDEYNNENLHFKFNQKYP 296
Db 210 MIEGKPLFPKGDHVFQFSIITDL---LGSP-----P 237

Qy 297 RKWDIILPCNDLDMK-----EIFTMIRYDRSKRITSK 331
Db 238 R--DVINTICSENTLKFVTSLPHRDPVPQFRFKEVTPDAVDLLERMLVDPKKRTAA 294

Qy 332 EIL 334
Db 295 DAL 297

RESULT 15
ID Q40482 PRELIMINARY; PRT; 294 AA.
AC Q40482;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-MAY-1999 (TEMBLrel. 10, Last annotation update)
DE CDC2.
GN CDC2.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
```


GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 1999, 14:48:40 ; Search time 11.2 seconds
(without alignments)
855.621 Million cell updates/sec

Title: US-09-072-994-14

Perfect score: 1791

Sequence: 1 MKLSYIDKELIYNSAISD.....IYDRSKRITSKEILQLMLD 339

Scoring table: BLOSUM62

Searched: 77977 seqs, 28268293 residues

Database : SwissProt_37.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	276.5	15.4	349	1 STYL_SCHPO	Q09892 schizosacch
2	273	15.2	296	1 CC2_DICDI	P34112 dictyosteli
3	271	14.9	368	1 CAK1_YEAST	P43568 saccharomyc
4	266	14.9	298	1 CDK2_MESAU	P48963 mesocricetu
5	265	14.8	297	1 CDK2_XENLA	P23437 xenopus lae
6	262.5	14.7	294	1 CC2_MAIZE	P23111 zea mays (m
7	262	14.6	298	1 CDK2_HUMAN	P49411 homo sapien
8	262	14.6	391	1 KC21_CHICK	P18868 gallus gall
9	262	14.6	391	1 KC21_HUMAN	P19138 homo sapien
10	262	14.6	391	1 KC21_RAT	P19139 rattus norv
11	261	14.6	298	1 CDK2_CARAU	P43450 carassius a
12	261	14.6	306	1 CSK1_SCHPO	P36615 schizosacch
13	259.5	14.5	311	1 CC23_TRYBB	P46666 trypanosoma
14	259	14.5	391	1 KC21_MOUSE	Q60737 mus musculu
15	259	14.5	350	1 KC22_XENLA	P28020 xenopus lae
16	258.5	14.4	294	1 CC21_ORYSA	P29618 oryza sativ
17	258	14.4	415	1 ERK1_DICDI	P42525 dictyosteli
18	258	14.4	391	1 KC21_RABIT	P33674 oryctolagus
19	256	14.3	305	1 CDK3_HUMAN	Q00526 homo sapien
20	255.5	14.3	297	1 CC2_SCHPO	P04551 schizosacch
21	255.5	14.3	435	1 HOG1_YEAST	P32485 saccharomyc
22	255.5	14.3	424	1 JNK2_HUMAN	P45984 homo sapien
23	255.5	14.3	420	1 KC2A_THEPA	P28547 theileria p
24	255	14.2	288	1 CC2H_PLAFK	Q07785 plasmodium
25	254.5	14.2	302	1 CC2_CARAU	P51958 carassius a
26	254.5	14.2	423	1 JNK2_RAT	P49186 rattus norv
27	251	14.0	350	1 KC22_CHICK	P21869 gallus gall
28	250.5	14.0	302	1 KC22_XENLA	P24033 xenopus lae
29	250	14.0	350	1 KC22_BOVIN	P20427 bos taurus
30	249	13.9	377	1 HOG1_CANAL	Q92207 candida alb
31	248.5	13.9	332	1 KC21_SCHPO	P40231 schizosacch
32	248	13.8	350	1 KC22_HUMAN	P19784 homo sapien
33	248	13.8	424	1 KC47_ORYSA	P29620 oryza sativ
34	246	13.7	317	1 CC28_CANAL	P43063 candida alb
35	246	13.7	361	1 MP38_XENLA	P47812 xenopus lae
36	245.5	13.7	302	1 CC21_XENLA	P35567 xenopus lae
37	245	13.7	333	1 KC22_ARATH	Q08466 arabidopsi
38	245	13.7	337	1 KC2A_DICDI	Q02720 dictyosteli
39	244.5	13.7	291	1 CC21_MEDSA	P24923 medicago sa
40	243.5	13.6	294	1 CC2A_ARATH	P24100 arabidopsi
41	243	13.6	387	1 MMK1_MEDSA	Q07176 medicago sa
42	242.5	13.5	294	1 CC2_VIGUN	P52389 vigna ungui
43	241	13.5	361	1 MP38_CYPCA	Q90336 cyprinus ca

44 240.5 13.4 303 1 CC2_CHICK P13863 gallus gall
45 240.5 13.4 297 1 CC2_HUMAN P06493 homo sapien

ALIGNMENTS

```
RESULT 1
STYL_SCHPO
ID STYL_SCHPO STANDARD; PRT; 349 AA.
AC Q09892;
DT 01-FEB-1996 (REL. 33, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE MITOGEN-ACTIVATED PROTEIN KINASE STYL (EC 2.7.1.-) (MAP KINASE STYL).
GN STYL OR SPCL OR SPAC24B11.06C.
OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; ARCHIASCOMYCETES;
OC SCHIZOSACCHAROMYCETALES; SCHIZOSACCHAROMYCETACEAE;
OC SCHIZOSACCHAROMYCES.
RN [1]
RP SEQUENCE FROM N.A.
RA MILLAR J.B.A., BUCK V., WILKINSON M.G.;
RL SUBMITTED (JUN-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
[2]
RN [2]
RP SEQUENCE FROM N.A.
RA SHIOZAKI K., RUSSELL P.;
RL SUBMITTED (OCT-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
[3]
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA ODELL C., CHURCHER C.M., BARRELL B.G., RAJANDREAM M.A., WALSH S.V.;
RL SUBMITTED (NOV-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- FUNCTION: INVOLVED OSMOREGULATION AND STRESS RESPONSE PATHWAYS.
CC -1- PROBABLY ACTIVATED BY WIS1.
CC -1- ENZYME REGULATION: ACTIVATED BY TYROSINE AND THREONINE
CC -1- PHOSPHORYLATION (BY SIMILARITY).
CC -1- PHOSPHORYLATION BELONGS TO THE CDC2/CDC28 SUBFAMILY OF SER/THR
CC -1- PROTEIN KINASES. STRONGEST SIMILARITY WITH OTHER MAP KINASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X89262; G897810; -
CC EMBL; U26739; G1022685; -
CC EMBL; Z67757; G1061294; -
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC FRAM; PF00069; pkinase; 1.
CC HSSP; Q16539; 1WFC.
CC TRANSFERASE; SERINE/THREONINE-PROTEIN KINASE; ATP-BINDING;
CC KW PHOSPHORYLATION.
CC DOMAIN 30 299 PROTEIN KINASE.
CC NP_BIND 26 34 ATP (BY SIMILARITY).
CC BINDING 49 49 ATP (BY SIMILARITY).
CC ACT_SITE 141 141 BY SIMILARITY.
CC MOD_RES 171 171 PHOSPHORYLATION (ACTIVATES THE KINASE)
CC (BY SIMILARITY).
CC MOD_RES 173 173 PHOSPHORYLATION (ACTIVATES THE KINASE)
CC (BY SIMILARITY).
CC SEQUENCE 349 AA; 40222 MW; B3385C56 CRC32;
SQ
```

Query Match 15.4%; Score 276.5; DB 1; Length 349;
Best Local Similarity 26.9%; Pred. No. 2.7e-10;
Matches 93; Conservative 60; Mismatches 114; Indels 79; Gaps 15;

DE	GN	KINASE).	199	IYKAPILGILITNVEYEIDWISGLIITGLYSENFOFVLVKDDKELTNDSHVSDYLLNQ	258
OS	XENOPUS	LAEVIS (AFRICAN CLAWED FROG).	167	WYRAPEILLOCKEYSTAVDIWSLGCFAEMTR	214
OC	EUKARYOTA;	METAZOA; CHORDATA; VERTEBRATA; AMPHIBIA; BATRACHIA; ANURA;			
OC	MESOBATRACHIA;	PIPOIDEA; PIPIDAE; XENOPODINAE; XENOPUS.			
RN	[1]		259	IFENFGTNPNTDFEELFCDEYNNENLHFKFNQYKPRKDWIILPRCNDLMLKEIFTK	318
RP	SEQUENCE FROM N.A.		215	IFRTGLTPDEVSWPGVTMPDY	265
RC	TISSUE=OVARY;				
RX	MEDLINE; 91126051.		319	MIRYDRSKRITSKEIL	334
RA	PARIS J., LE GUELLEC R., COUTURIER A., LE GUELLEC K., OMILLI F.,				
RA	CAMONIS J., MACNETT S., PHILIPPE M.,		266	MLOYDSNKRISAKVAL	281
RT	"Cloning by differential screening of a xenopus cDNA coding for a				
RT	protein highly homologous to cdc2".				
RL	PROC. NATL. ACAD. SCI. U.S.A. 88:1039-1043(1991).				
RN	[2]				
RP	PHOSPHORYLATION AT THR-160.				
RX	MEDLINE; 93345457.				
RA	POON R.Y.C., YAMASHITA K., ADAMCZEWSKI J.P., HUNT T., SHUTTLEWORTH J.;				
RT	"The cdc2-related protein p40M015 is the catalytic subunit of a				
RT	protein kinase that can activate p33cdc2 and p34cdc2".				
RL	EMBO J. 12:3123-3132(1993).				
CC	-1- FUNCTION: PROBABLY INVOLVED IN THE CONTROL OF THE CELL CYCLE.				
CC	INTERACTS WITH CYCLINS A, D, OR E.				
CC	-1- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES				
CC	THE ENZYME, WHILE PHOSPHORYLATION AT THR-160 ACTIVATES IT (BY				
CC	SIMILARITY).				
CC	-1- DEVELOPMENTAL STAGE: SYNTHESIZED IN UNFERTILIZED EGG, BUT NO				
CC	LONGER MADE IN THE EARLY EMBRYO.				
CC	-1- SIMILARITY: WITH THE CONSERVED CATALYTIC DOMAINS OF SER/THR-				
CC	PROTEIN KINASES. BELONGS TO THE CDC2/CDKX SUBFAMILY.				
CC					
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC					
CC	EMBL; X14227; 654666; -				
DR	PIR; A37871; A37871.				
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.				
DR	PROSITE; PS00108; PROTEIN_KINASE_ST; 1.				
DR	PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.				
DR	PFAM; PF00069; pkinase; 1.				
DR	HSP; P24941; 1HCK.				
DR	TRANSFERASE; SERINE/THREONINE-PROTEIN KINASE; ATP-BINDING;				
KW	CELL CYCLE; CELL DIVISION; MITOSIS; PHOSPHORYLATION.				
FT	DOMAIN 4 286				
FT	NP_BIND 10 18				
FT	BINDING 33 33				
FT	ACT_SITE 127 127				
FT	MOD_RES 14 14				
FT	MOD_RES 15 15				
FT	MOD_RES 160 160				
FT	SEQUENCE 297 AA; 33969 MW; 8ACFB99B CRC32;				
QY	Query Match	14.8%; Score 265; DB 1; Length 297;			
Db	Best Local Similarity	26.6%; Pred. No. 1.1e-09;			
Db	Matches	84; Conservative 57; Mismatches 121; Indels 54; Gaps 11;			
QY	21	LYTATDKFNLPVCLKIV--DEDSLPHSIRHEIFILKTLKPHNIEYFNDLKIVDDV	78		
Db	18	VKARNRETGEIVALKKRLTETEGVPSTAIRISLLKELN--HPNIVKLLDVHTENKL	76		
QY	79	ILVTKLYRDLSQLIEITKYCKRTFRFYINGNLVSNQYTLANEIEEKDKLWLSKMS	138		
Db	77	YLVEEFLNODLK-----KFMDRSN--ISGISLALV-----KSYLFOLLQ	113		
QY	139	GLEFTHSGIHRDKPSNIFARDITQPIIGDPICVDLKLPPKDPMPMAKVIDYSTG	198		
Db	114	GLAFCHSHRVLRDLKPNLINSDAIK--LADFLGARAFGVPRFETH-----EVVTL	166		

Query Match 14.7%; Score 262.5; DB 1; Length 294;
Best Local Similarity 24.9%; Pred. No. 1.5e-09;
Matches 86; Conservative 59; Mismatches 123; Indels 77; Gaps 13;

QY 5 DYYIDKELIYNSAISIYTAIDKFNMLPVCLK---IVDEDFSLPHSHREIFILKTLKP 61
DB 2 EYKVEKIGEGYGVVYKALDRAITETALKKIRLEQDEGVSTAI-REISLLKEMN- 60
QY 62 HPIIIEYFNLDKIYDDVILYTKLYRDLDSOLIE-ITKYCKRTRTRIYGINGNLVSNQYTL 120
DB 60 HGNIVRLHDVVHSEKRIYLVFEYLDLKKEMDSCPEFAKNPTL----- 104
QY 121 ANEIEEDKIKLWKSMSGLEFTHSGIIRHDIKPSNFFARDDITQPI-IGDFDICYDL 179
DB 104 -----IKSYLVQILHGVAYCHSHRVLRDLKPNQLI--DRRTNALKLADFLARAF 153
QY 180 KLPRDEPPMAKYIDVSTGIYKAPEILGITNYEYEDISWISGILITGLYSENFQSVLVK 239
DB 154 GIPVTRTFTH-----EVTLYWRAPEILLGARQYSTPVDVWSVGCIF-----AEMVN 199
QY 240 DDKELTNDVSHVSDIYLLNQIFENFGTPN-----LTFDEFELEFCDYNNENLHFKKF 290
DB 200 QKPLFPDGEIDELF---KIFRLITGTPNEQSWPGVSCULPFF-----KT 239
QY 291 NLQYPRKDWDIILPRCNDLDMKEIFTRMIRYDRSKRITSKEILQ 335
DB 240 APRWQADLATVVPNL-DPAGLDLLSKMLRYEPSKRITARQALE 283

RESULT 7
CDK2_HUMAN
ID CDK2_HUMAN STANDARD; PRT; 298 AA.
AC P24941;
DT 01-MAR-1992 (REL. 21, CREATED)
DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE CELL DIVISION PROTEIN KINASE 2 (EC 2.7.1.-) (P33 PROTEIN KINASE).
GN CDK2.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 91330891.
RA ELLEDGE S.J., SPOTSWOOD M.R.;
RT "A new human p34 protein kinase, CDK2, identified by complementation
RT of a cdc28 mutation in *Saccharomyces cerevisiae*, is a homolog of
RT *Xenopus* Egi.";
RL EMBO J. 10:2653-2659(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 91367262.
RA TSAI L.-H., HARLOW E., MEYERSON M.;
RT "Isolation of the human cdk2 gene that encodes the cyclin A- and
RT adenovirus E1A-associated p33 kinase.";
RL NATURE 353:174-177(1991).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE; 92020980.
RA NINOMIYA-TSUJI J., NOMOTO S., YASUDA H., REED S.I., MATSUMOTO K.;
RT "Cloning of a human cDNA encoding a CDC2-related kinase by
RT complementation of a budding yeast cdc28 mutation.";
RL PROC. NATL. ACAD. SCI. U.S.A. 88:9006-9010(1991).
RN [4]
RP PHOSPHORYLATION SITES.
RX MEDLINE; 93010995.
RA GU Y., ROSENBLATT J., O'MORGAN D.O.;
RT "Cell cycle regulation of CDK2 activity by phosphorylation of Thr160
RT and Tyr15.";
RL EMBO J. 11:3995-4005(1992).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
RX MEDLINE; 93288132.
RA DE BONDY H.L., ROSENBLATT J., JANCARIK J., JONES H.D.,
RA MORGAN D.O., KIM S.-H.;
RT "Crystal structure of cyclin-dependent kinase 2.";

RL NATURE 363:595-602(1993).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF COMPLEX WITH CYCLIN A.
RX MEDLINE; 93535811.
RA JEFFREY P.D., RUSSO A.A., POLYAK K., GIBBS E., HURWITZ J.,
RA MASSAGUE J., PAVLETICH N.P.;
RT "Mechanism of CDK activation revealed by the structure of a
RT cyclinA-CDK2 complex.";
RL NATURE 376:313-320(1995).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.33 ANGSTROMS) OF COMPLEX WITH L868276.
RX MEDLINE; 96184476.
RA DE AZEVEDO W.F. JR., MULEER-DIECKMANN H.-J., SCHULZE-GAHMEN U.,
RA WOLAND P.J., SAUSVILLE E., KIM S.-H.;
RT "Structural basis for specificity and potency of a flavonoid
RT inhibitor of human CDK2, a cell cycle kinase.";
RL PROC. NATL. ACAD. SCI. U.S.A. 93:2735-2740(1996).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF COMPLEX WITH CG2A AND KIP1.
RX MEDLINE; 96300318.
RA RUSSO A.A., JEFFREY P.D., PATTEN A.K., MASSAGUE J., PAVLETICH N.P.;
RT "Crystal structure of the p27kip1 cyclin-dependent-kinase inhibitor
RT bound to the cyclin A-Cdk2 complex.";
RL NATURE 382:325-331(1996).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF COMPLEX WITH CG2A.
RX MEDLINE; 96313126.
RA RUSSO A.A., JEFFREY P.D., PAVLETICH N.P.;
RT "Structural basis of cyclin-dependent kinase activation by
RT phosphorylation.";
RL NAT. STRUCT. BIOL. 3:696-700(1996).
RN [10]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX MEDLINE; 97075215.
RA SCHULZE-GAHMEN U., DE BONDY H.L., KIM S.-H.;
RT "High-resolution crystal structures of human cyclin-dependent kinase
RT 2 with and without ATP: bound waters and natural ligand as guides for
RT inhibitor design.";
RL J. MED. CHEM. 39:4540-4546(1996).
RN [11]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE; 97475219.
RA LAWRIE A.M., NOBLE M.E.M., TUNNAH P., BROWN N.R., JOHNSON L.N.,
RA ENDICOTT J.A.;
RT "Protein kinase inhibition by staurosporine revealed in details of
RT the molecular interaction with CDK2.";
RL NAT. STRUCT. BIOL. 4:796-801(1997).
RN [12]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF COMPLEX WITH CKS1.
RX MEDLINE; 96182647.
RA BOURNE Y., WATSON M.H., HICKEY M.J., HOLMES W., ROCQUE W., REED S.I.,
RA TAINER J.A.;
RT "Crystal structure and mutational analysis of the human CDK2 kinase
RT complex with cell cycle-regulatory protein CksHs1.";
RL CELL 84:863-874(1996).
RN [13]
RP FUNCTION: PROBABLY INVOLVED IN THE CONTROL OF THE CELL CYCLE.
CC INTERACTS WITH CYCLIN A, D, OR E. ACTIVITY OF CDK2 IS MAXIMAL
CC DURING S PHASE AND G2.
CC ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES
CC THE ENZYME, WHILE PHOSPHORYLATION AT THR-160 ACTIVATES IT.
CC SIMILARITY: WITH THE CONSERVED CATALYTIC DOMAINS OF SER/THR-
CC PROTEIN KINASES. BELONGS TO THE CDC2/CDK SUBFAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X61622; G29849; -.
DR EMBL; X62071; G312803; -.

```
DR EMBL: M68520; G180178; -.
DR PIR: A41227; A41227.
DR PIR: S16520; S16520.
DR PIR: S17873; S17873.
DR PDB: 1JIN; 27-JAN-97.
DR PDB: 1HCK; 07-DEC-96.
DR PDB: 1HCL; 07-DEC-96.
DR PDB: 1AQL; 12-NOV-97.
DR PDB: 1JST; 11-JAN-97.
DR PDB: 1JSU; 29-JUL-97.
DR PDB: 1BUH; 03-SEP-98.
DR PDB: 1B38; 23-DEC-98.
DR PDB: 1B39; 23-DEC-98.
DR MIM: 116953; -.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PFAM: PF00069; pkinase; 1.
DR TRANSFERRASE; SERINE/THREONINE-PROTEIN KINASE; ATP-BINDING;
KW CELL CYCLE; CELL DIVISION; MITOSIS; PHOSPHORYLATION; 3D-STRUCTURE.
KW DOMAIN 4 286 PROTEIN KINASE.
FT NP_BIND 10 18 ATP (BY SIMILARITY).
FT BINDING 33 33 ATP (BY SIMILARITY).
FT ACT_SITE 127 127 BY SIMILARITY.
FT MOD_RES 14 14 PHOSPHORYLATION.
FT MOD_RES 15 15 PHOSPHORYLATION.
FT MOD_RES 160 160 PHOSPHORYLATION (BY CAK).
FT MUTAGEN 14 14 T->A: INCREASE ACTIVITY 2 FOLD.
FT MUTAGEN 15 15 Y->F: INCREASE ACTIVITY 2 FOLD.
FT MUTAGEN 160 160 T->A: ABOLISHES ACTIVITY.
KW SEQUENCE 298 AA; 33929 MW; 7538D6FA CRC32;
Query Match 14.6%; Score 262; DB 1; Length 298;
Best Local Similarity 26.2%; Pred. No. 1.7e-09;
Matches 83; Conservative 59; Mismatches 119; Indels 56; Gaps 11;
QY 21 IYTAIDKNNLPVCLKIV--DEFSIPLPHSIHREIFILKTLKPHNIIYEYNDKIYDDV 78
DB 18 VYKARKNTIGVWALKIRLDTETEGVSTAIRESLSKELN-HPNIVKLLDVHTHNTKL 76
QY 79 ILVTKLYRDLQSLEIETKYCKTRTFYINGNLVSNQYTLANEIEKDKLWKLKSMSS 138
DB 77 VLVPEELHODLKKFMDASA-----LUGLPLEP-----IKSYLFOLLQ 113
QY 139 GLEFTHSGIITHRDKSNIFFARDDITQPIGDFDICYDLKLPKDPMPAKYI-DYST 197
DB 114 GLAFCHSHRVJHRLDKPONLLINTEGAIR--LADFLARAFGV-----PVRTTHEVVT 165
QY 198 GIYKAPELILGITNYEYIDWSLGIILGLXSENFSQVVLVKKDELNDSHVSDLYLN 257
DB 166 LWYRAPEILLGCKYYSTAVDWSLGCIF-----ADMVTRRALFPGDSEIDQLF--- 214
QY 258 OIFENFGTPNLTDFEDELFCDEYNENLHFKKFNQKYPKDWDLILPRCNDLDMKEIFT 317
DB 214 RIFRLTGPDEVWPGVTSMPDI-----KPSPPKWARQDFSKVVPPLDED-GRSLLS 264
QY 318 KMIRYDRSKRITSKEIL 334
DB 265 QMLHYDPNKRISAKAAL 281
RESULT 8
ID KC21_CHICK STANDARD; PRT; 391 AA.
AC P21868;
DT 01-MAY-1991 (REL. 18, CREATED)
DT 01-MAY-1991 (REL. 18, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE CASIN KINASE II, ALPHA CHAIN (CK II) (EC 2.7.1.37).
OS GALLUS GALLUS (CHICKEN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ARCHOSAURIA; AVES;
OC NEOGNATHAE; GALLIFORMES; PHASIANIDAE; PHASIANINAE; GALLUS.
```

```
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 91115855.
RA MARIDOR G., PARK W., KREK W., NIGG E.A.;
RT "Casein Kinase II. cDNA sequences, developmental expression, and
RT tissue distribution of mRNAs for alpha, alpha", and beta subunits of
RT the chicken enzyme.";
RL J. BIOL. CHEM. 266:2362-2368(1991).
CC -1- FUNCTION: CASIN KINASES ARE OPERATIONALLY DEFINED BY THEIR
CC AS SUBSTRATES
CC PREFERENTIAL UTILIZATION OF ACIDIC PROTEINS SUCH AS CASEINS
CC -1- FUNCTION: THE ALPHA AND ALPHA' CHAINS CONTAIN THE CATALYTIC SITE.
CC -1- SUBUNIT: TETRAMER COMPOSED OF AN ALPHA CHAIN, AN ALPHA' AND TWO
CC BETA CHAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: M59456; G211534; -.
CC PIR: A38611; A38611.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
CC PFAM: PF00069; pkinase; 1.
CC HSP; P28523; IAG0.
CC TRANSFERRASE; SERINE/THREONINE-PROTEIN KINASE; ATP-BINDING.
KW DOMAIN 39 324 PROTEIN KINASE.
FT NP_BIND 45 53 ATP (BY SIMILARITY).
FT BINDING 68 68 ATP (BY SIMILARITY).
FT ACT_SITE 156 156 BY SIMILARITY.
FT SEQUENCE 391 AA; 45190 MW; EDB98195 CRC32;
Query Match 14.6%; Score 262; DB 1; Length 391;
Best Local Similarity 26.5%; Pred. No. 2.3e-09;
Matches 90; Conservative 59; Mismatches 125; Indels 66; Gaps 13;
QY 5 DYITDKELIYNSAISDIYTAIDKNNLPVCLKIVDEFSIPLPHSIHREIFILKTLKPHN 64
DB 38 DYQLVRKL-GRGKYSVFEEAINITNKKVVKILK---PVKKKKIKREIKILENLGGPN 93
QY 65 IIEYFNDLKIYDDVILVTKLYRDLQSLEIETKYCKTRTFYINGNLVSNQYTLANEI 124
DB 94 IITLADIVK--DVP-----SRTPALVFEHVNTDFKO--LYQTL 128
QY 125 EEDKIKLWKLKSMSSGLEFTHSGIITHRDKSNIFFARDDITQPIGDFDICYDLKLPK 184
DB 129 TDVDIRFYVYELKALDYCHSGIMHREDVKPHNVIDHEHKLRLI-----DWGLAEF 181
QY 185 DEPPMAKYIDVSTGYKAPELILGITNYEYIDWSLGIILGLY--SENFOSVLVYDKD 242
DB 182 YHFGQEVNVRVASRFGKPELLDYQMDYDLSLDMWSLGLASIMIFRKEPF----- 233
QY 243 ELTNDSHVSDLY--LLNQIFENFGTPNLTDFEDELFCDEYNENLHFKKFN--LQYPRKD 299
DB 233 -----FHGHNDYDQLVRIAKVLGTEDLYD-----YIDKNIE--LDPRFNDILGRSKR 280
QY 300 WDILPRCNDLDMK-----EFTKMIRYDRSKRITSKEILO 335
DB 281 WERFVHSEHQHLVSPALDFDLKLLRYDHQSLRTAREAME 320
RESULT 9
ID KC21_HUMAN STANDARD; PRT; 391 AA.
AC P19138; P20426;
DT 01-NOV-1990 (REL. 16, CREATED)
DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
```

```

DR 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE CASEIN KINASE II, ALPHA CHAIN (CK II) (EC 2.7.1.37).
GN CSNK2A1 OR CK2A1.
OS HOMO SAPIENS (HUMAN), AND BOS TAURUS (BOVINE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=HUMAN;
RX MEDLINE; 89323123.
RA MEISNER H., HELLER-HARRISON R., BUXTON J., CZECH M.P.;
RL "Molecular cloning of the human casein kinase II alpha subunit.";
RL BIOCHEMISTRY 28:4072-4076(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=HUMAN;
RX MEDLINE; 91070071.
RA LOZEMAN F.J., LITCHFIELD D.W., PIENING C., TAKIO K., WALSH K.A.,
RA KREBS E.G.;
RL "Isolation and characterization of human cDNA clones encoding the
RL alpha and the alpha' subunits of casein kinase II.";
RL BIOCHEMISTRY 29:8436-8447(1990).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=BOVINE;
RX MEDLINE; 93223703.
RA OLE-MOIVOI O.K., BROWN W.C., IAMS K.P., NAYAR A., TSUKAMOTO T.,
RA MACKLIN M.D.;
RL "Evidence for the induction of casein kinase II in bovine lymphocytes
RL transformed by the intracellular protozoan parasite Theileria
RL parva.";
RL EMBO J. 12:1621-1631(1993).
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES=BOVINE; TISSUE=TESTIS;
RA WATANABE M., YUGE M., MAEDA O., OHNO S., KAWASAKI H., SUZUKI K.,
RA HIDAKA H.;
RL SUBMITTED (OCT-1990) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [5]
RP SEQUENCE OF 50-59; 103-122; 230-239; 248-279 AND 284-303.
RC SPECIES=BOVINE; TISSUE=TESTIS;
RX MEDLINE; 90237072.
RA LITCHFIELD D.W., LOZEMAN F.J., PIENING C., SOMMERCOORN J., TAKIO K.,
RA WALSH K.A., KREBS E.G.;
RL "Subunit structure of casein kinase II from bovine testis.
RL Demonstration that the alpha and alpha' subunits are distinct
RL polypeptides.";
RL J. BIOL. CHEM. 265:7638-7644(1990).
CC -1- FUNCTION: CASEIN KINASES ARE OPERATIONALLY DEFINED BY THEIR
CC PREFERENTIAL UTILIZATION OF ACIDIC PROTEINS SUCH AS CASEINS
CC AS SUBSTRATES.
CC -1- FUNCTION: THE ALPHA AND ALPHA' CHAINS CONTAIN THE CATALYTIC SITE.
CC -1- SUBUNIT: TETRAMER COMPOSED OF AN ALPHA CHAIN, AN ALPHA' AND TWO
CC BETA CHAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; J02853; G598147; -
DR EMBL; M52265; G177994; -
DR EMBL; M93665; G162777; -
DR EMBL; X54962; G611; -
DR PIR; A30319; A30319.
DR PIR; A35838; A35838.
DR PIR; S21335; S21335.
DR PIR; A35206; A35206.
DR MIM; I15440; -
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.

```

```

DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PFAM; PF00069; pkinase; 1.
DR HSSP; P28523; IAGO.
KW TRANSFERASE; SERINE/THREONINE-PROTEIN KINASE; ATP-BINDING.
FT DOMAIN; 39 324 PROTEIN KINASE.
FT NP_BIND; 45 53 ATP (BY SIMILARITY).
FT BINDING; 68 68 ATP (BY SIMILARITY).
FT ACT_SITE; 156 156 BY SIMILARITY.
SQ SEQUENCE 391 AA; 45143 MW; 559AA2A0 CRC32;

Query Match 14.6%; Score 262; DB 1; Length 391;
Best Local Similarity 26.5%; Pred. No. 2.3e-09;
Matches 90; Conservative 59; Mismatches 125; Indels 66; Gaps 13;

QY 5 DYIDKELYNISAIIDYTAIDKFNLPVCLIKVIDEFLSPPHSHREIFILTKLPHNP 64
DB 38 DYQLVRKL-GRGKYSVFEAINITNNEKVVVKILK---PVKKKKIKREIKILENLRGSPN 93
QY 65 IIEYFNDLKIYDDVILVTLYKLYRDLSQLIEITKYCKRTRFYINGNLVSNQYTLANEI 124
DB 94 IITLADIVK-DPV-----SRTPALVFEHVNTDFKQ--LYQTL 128
QY 125 EEKDIKLWLKSMSSGLEFIHSGIIRHDIKPSIFFAFDDITQPIIGDFDICYDLKLPKP 184
DB 129 TDYDIRFYWEILKALDYCHSGIMHROVKPHNVIMDEHHRKRLI-----DWGLAEF 181
QY 185 DEPMKAYIDVSTGYKAPELILGITNTEYEYEDISLGIITGLY--SENFSVLVVDK 242
DB 182 YHPGQSVINRVASRYFKGPELLVDYQMDYSLDMWSLGLCMASLMIFRKEPF----- 233
QY 243 ELTNDSHVSDLY-LLNQIFENFGTNPNTLDFEDELFCDDYNNENLHFKFN--LOKYPKRD 299
DB 233 -----FHGHNDVQLVRIAKVLGTEDLYD-----YIDKNIE--LDPRNDILGHSRKR 280
QY 300 WDILPRCNDLMLK-----EFTKMYRDSKRITSKEILQ 335
DB 281 WERFVHSEHQHLVSPALDFLDKLLRYDHQSRLTAREAME 320

RESULT 10
KC21_RAT ID KC21_RAT STANDARD; PRT: 391 AA.
AC P19139;
DT 01-NOV-1990 (REL. 16, CREATED)
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE CASEIN KINASE II, ALPHA CHAIN (CK II) (EC 2.7.1.37).
GN CSNK2A1.
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER.
RA AHMED K., DAVIS A., HANTEN J., LAMBERT D., MCIVOR R.S.,
RA GUELI S.A.;
RL SUBMITTED (XXX-1993) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE OF 8-391 FROM N.A.
RX MEDLINE; 89323123.
RA MEISNER H., HELLER-HARRISON R., BUXTON J., CZECH M.P.;
RL "Molecular cloning of the human casein kinase II alpha subunit.";
RL BIOCHEMISTRY 28:4072-4076(1989).
CC -1- FUNCTION: CASEIN KINASES ARE OPERATIONALLY DEFINED BY THEIR
CC PREFERENTIAL UTILIZATION OF ACIDIC PROTEINS SUCH AS CASEINS
CC AS SUBSTRATES.
CC -1- FUNCTION: THE ALPHA AND ALPHA' CHAINS CONTAIN THE CATALYTIC SITE.
CC -1- SUBUNIT: TETRAMER COMPOSED OF AN ALPHA CHAIN, AN ALPHA' AND TWO
CC BETA CHAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```



```

GN CSK1 OR SPACID4.06C.
OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; ARCHIASCOMYCETES;
OC SCHIZOSACCHAROMYCETALES; SCHIZOSACCHAROMYCETACEAE;
OC SCHIZOSACCHAROMYCES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 93223713.
RA MOLZ L., BEACH D.;
RT "Characterization of the fission yeast mcs2 cyclin and its associated
RL protein kinase activity";
RL EMBO J. 12:1723-1732(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA LYE G., CHURCHER C.M., BARRELL B.G., RAJANDREAM M.A., WALSH S.V.;
RL SUBMITTED (FEB-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- FUNCTION: NOT KNOWN. COULD FUNCTION IN ASSOCIATION WITH CYCLIN
CC MCS2.
CC -!- SIMILARITY: WITH THE CONSERVED CATALYTIC DOMAINS OF SER/THR-
CC PROTEIN KINASES. BELONGS TO THE CDC2/CDKX SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S59896; G299549; -
DR EMBL; Z69239; E1188809; -
DR PIR; S39151; S39151.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PFAM; PF00069; pkinase; 1.
DR HSSP; P24941; 1AQL.
DR TRANSFERASE; SERINE/THREONINE-PROTEIN KINASE; ATP-BINDING.
KW DOMAIN 11 306 PROTEIN KINASE.
FT NP_BIND 17 25 ATP (BY SIMILARITY).
FT BINDING 40 40 ATP (BY SIMILARITY).
FT ACT_SITE 129 129 BY SIMILARITY.
FT SEQUENCE 306 AA; 34672 MW; 514DE9CE CRC32;
Query Match 14.6%; Score 261; DB 1; Length 306;
Best Local Similarity 25.1%; Pred. No. 2e-09;
Matches 88; Conservative 62; Mismatches 127; Indels 74; Gaps 15;
QY 1 MKLSDYI-----DKELIYNSAIDYITADKFNLPVCLKIVDEDFSLPPHSIHREIFIL 56
DB 1 MKSVGHFVPLWTDIRHLTDGTITSEFVGVGRKNSKKLYIKVQGLVFKRPDAMRGKLI 60
QY 57 KTL-KPH-PNITIEYNDLKIYDDVILVNLKLYDYLSQ-----LEIKYCKRTRIFYING 111
DB 61 ESIGHPIHRIYVDSIDNEA-GSVLTITFSKFSVLSDVNDMEISIDTKCKIVLQ----- 113
QY 112 NLVSNQYTLANEIEEKDKIKLWKSSSGLEFTHSQGIHHRDKPSNFIFFARDITQP-II 170
DB 113 -----ISSALEYLEKHGILHROIHPNNIL--DSMNGPAYL 146
QY 171 GDFDICYDLKLPKPPMAKVIDVSTGIYKAPELILGITNYEYIDTWSLGIILTGLYS 230
DB 147 SDFSIAWSKQHPGEEVQELIP--QIGTGHYRAIETLFGCHSYGHEVDRTWTFGLIAELFS 204
QY 231 ENFQSVLVKDKELNDSHVSLYLLNOLFENFGPN-----LTDPELFCDEYNEN 284
DB 205 NQ-----ALFDDG--SSEGWPSRLRTSSIIQLGTPNPMPELSTFPD----- 248
QY 285 LHKFNKLOKYPKRDWDIILPCNDDLMKEIFTMKIRYDRSKRITSKEILQ 335
DB 248 --WNKFIFHEYPKPSWELPSV-DTSLQIVSHLVTV--SNRASPSVIE 293

```

```

RESULT 13
CC23_TRYBB ID CC23_TRYBB STANDARD; PRT; 311 AA.
AC P54666;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE CELL DIVISION CONTROL PROTEIN 2 HOMOLOG 3 (EC 2.7.1.-).
GN CRK3.
OS TRYPANOSOMA BRUCEI BRUCEI.
OC EUKARYOTA; EUKLENZOEA; KINETOPLASTIDA; TRYPANOSOMATIDAE; TRYPANOSOMA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ISTAT;
RX MEDLINE; 96009893.
RA MOTTRAM J., SMITH G.;
RT "A family of trypanosome cdc2-related protein kinases.";
RL GENE 162:147-152(1995).
CC -!- FUNCTION: PROBABLY INVOLVED IN THE CONTROL OF THE CELL CYCLE.
CC -!- ENZYME REGULATION: PHOSPHORYLATION AT SER-33 OR TYR-34 INACTIVATES
CC THE ENZYME.
CC -!- SUBUNIT: FORMS A STABLE BUT NON-COVALENT COMPLEX WITH A REGULATORY
CC SUBUNIT AND WITH A CYCLIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CDC2/CDC28 SUBFAMILY OF SER/THR
CC PROTEIN KINASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X74617; G397365; -
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PFAM; PF00069; pkinase; 1.
DR HSSP; P24941; 1AQL.
DR TRANSFERASE; SERINE/THREONINE-PROTEIN KINASE; ATP-BINDING;
KW PHOSPHORYLATION.
FT DOMAIN 23 306 PROTEIN KINASE.
FT NP_BIND 29 37 ATP (BY SIMILARITY).
FT BINDING 52 52 ATP (BY SIMILARITY).
FT ACT_SITE 145 145 BY SIMILARITY.
FT MOD_RES 33 33 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 34 34 PHOSPHORYLATION (BY SIMILARITY).
FT SEQUENCE 311 AA; 35047 MW; 6594D66E CRC32;
Query Match 14.5%; Score 259.5; DB 1; Length 311;
Best Local Similarity 25.9%; Pred. No. 2.4e-09;
Matches 91; Conservative 57; Mismatches 113; Indels 91; Gaps 15;
QY 5 DYVIDKELIYNSAIDYITADKFNLPVCLKIV-----DEDFSLPPHSIHREIFILKTL 59
DB 21 DRNRMDILLEGTYGVYRAVDRATGQIVALKKRVLDRTDEGI---PQALREVSLQEI 77
QY 60 KPHNIIIEYFN-----DLKIYDDVILVNLKLYRYDLSOLIEITKYCKRTRIFYINGNLSV 115
DB 78 H-HPNIVNLLDVICADGKLY---LIFEYVDHDLKALE-----KRGAF----- 118
QY 116 NQYTLANEIEEKDKIKLWKSSSGLEFTHSQGIHHRDKPSNFIFFARDITQPIIGDFDI 175
DB 118 -----TGTTLKIIYQLLEGISFCFCHRRIVHRDLKAPANILVTDNSVK--IADFGL 166
QY 176 CYDLKLPKPPMAKVI-DVSTGIYKAPELILGITNYEYIDTWSLGIILTGLYSENFQ 234
DB 167 ARAFOI-----PMHTYTHEVTLNTRAPEILLGKHKHYTPAVDMWSIGCIFAEAR--G 217

```

```

Db      129 TDYDIRFYMEILLKALDYCHSGMIMHRDVKPHNVMIDHEHRKURLI-----DWGLAEF   181
Qy      185 DEPPMAKYDVSTGIVKAPELLIGINYEYEDWSLGIILTCLY--SENFQSVLVKKDK   242
Db      182 YHFGQEYNVRVSRKYFKPELLVDYQMYDSLDWMSLGCMASIFRKEPF-----233
Qy      243 ELTNDSHVSPLY-LLNQIFENFTGNLTDFEELFCDEVNNELHKFFN--LQKYPRKD   299
Db      233 -----FHGHNYDLVRIAKVLGTEDLYD-----YIDKYNI--LDPRFNDILGRHSKR   280
Qy      300 WDILPCRDLMK----EFTMKIRYDRSKRTSKEILQ   335
Db      281 WERFVHSENHLVSPEALDFDLKLRYDHQSRLTVREAME   320

RESULT 15
KC22_XENLA
ID      KC22_XENLA          STANDARD;          PRT;          350 AA.
AC      P28020;
DT      01-AUG-1992 (REL. 23, CREATED)
DT      01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
DT      01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE      CASEIN KINASE II, ALPHA' CHAIN (CK II) (EC 2.7.1.37).
CS      XENOPUS LAEVIS (AFRICAN CLAWED FROG).
CC      EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; AMPHIBIA; BATRACHIA; ANURA;
OC      MESOBATRACHIA; PIPOIDEA; PIPIDAE; XENOPODINAE; XENOPUS.
[1]
FN      SEQUENCE FROM N.A.
EC      TISSUE-Ovary.
FX      MEDLINE; 92183811.
FA      JEDLIACKI A., HINRICHS M.V., ALLENDE C., ALLENDE J.E.;
FT      "The cDNAs coding for the alpha- and beta-subunits of xenopus laevis
FT      casein kinase II.";
FL      FEBS LETT. 297:280-284(1992).
CC      -!- FUNCTION: CASEIN KINASES ARE OPERATIONALLY DEFINED BY THEIR
CC      PREFERENTIAL UTILIZATION OF ACIDIC PROTEINS SUCH AS CASEINS
CC      AS SUBSTRATES.
CC      -!- FUNCTION: THE ALPHA AND ALPHA' CHAINS CONTAIN THE CATALYTIC SITE.
CC      -!- SUBUNIT: Tetramer composed of an alpha chain, an alpha' and two
CC      beta chains.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation at
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; X62375; G64628; -.
DR      PIR; S18897; S18897.
DR      PIR; S20404; S20404.
DR      PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR      PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR      PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR      PFAM; PF00069; pkinase; 1.
DR      HSP; P28523; 1A6O.
KW      TRANSFERASE; SERINE/THREONINE-PROTEIN KINASE; ATP-BINDING.
FT      DOMAIN              39..324    PROTEIN KINASE.
FT      NP_BIND             45..53     ATP (BY SIMILARITY).
FT      BINDING            68..68     ATP (BY SIMILARITY).
FT      ACT_SITE           156..156    BY SIMILARITY.
SQ      SEQUENCE          350 AA; 41454 MW; 8956ED02 CRC32;

Query Match          14.5%; Score 259; DB 1; Length 350;
Best Local Similarity 26.5%; Fred. No. 3e-09;
Matches          90; Conservative          58; Mismatches          125; Indels          66; Gaps          13;

Qy      5 DYVIDKELYNSAIGDIYTAIDKFNNLPVCLIKVIDEFDPSLPHPHSIHREIFILTKTLKPHNP   64
          |||:|||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
          29 DYVIDKELYNT-CPGVCFVSVEFAINTNTNEKVKKVLIK---PVKKKKIKIKREIKILENLRGGP   93

```

Search completed: November 6, 1999, 00:45:07
Job time: 35787 sec

Result No.	§				ID	Description
	Score	Query Match	Length	DB		
1	261	14.6	298	2	US-08-874-347-25	Sequence 25, Appl
2	258.5	14.4	294	2	US-08-874-347-25	Sequence 26, Appl
3	255.5	14.3	297	2	US-08-874-347-22	Sequence 22, Appl
4	255.5	14.3	424	3	PCR-US94-12913A-18	Sequence 18, Appl
5	251.5	14.0	365	1	US-08-674-612-2	Sequence 2, Appl
6	251.5	14.0	365	2	US-08-920-296-2	Sequence 2, Appl
7	246.5	13.8	365	2	US-08-746-788-2	Sequence 2, Appl
8	246	13.7	317	2	US-08-463-090B-9	Sequence 9, Appl
9	246	13.7	317	2	US-08-874-347-18	Sequence 18, Appl
10	240.5	13.4	297	2	US-08-874-347-24	Sequence 24, Appl
11	238.5	13.3	297	2	US-08-874-347-23	Sequence 23, Appl
12	237.5	13.3	297	1	US-08-176-620A-16	Sequence 16, Appl
13	237.5	13.3	297	2	US-08-461-985-16	Sequence 16, Appl
14	236.5	13.2	353	1	US-08-176-620A-14	Sequence 14, Appl
15	236.5	13.2	353	2	US-08-461-985-14	Sequence 14, Appl
16	236.5	13.2	384	3	PCR-US94-08119-12	Sequence 12, Appl
17	236.5	13.2	384	3	PCR-US94-08119-12	Sequence 12, Appl
18	235.5	13.1	323	2	PCR-US94-12913A-12	Sequence 12, Appl
19	235	13.1	360	2	US-08-874-347-21	Sequence 21, Appl
20	235	13.1	360	2	US-08-469-421-14	Sequence 14, Appl
21	235	13.1	360	2	US-08-250-975-14	Sequence 14, Appl
22	235	13.1	360	2	US-08-605-002A-14	Sequence 14, Appl
23	235	13.1	360	2	US-08-950-449A-14	Sequence 14, Appl
24	233	13.0	364	3	PCR-US94-10529-14	Sequence 14, Appl
25	232	13.0	384	2	US-08-651-940-2	Sequence 2, Appl
26	232	13.0	282	2	US-08-318-947A-19	Sequence 19, Appl
27	232	13.0	274	2	US-08-318-947A-20	Sequence 20, Appl
28	232	13.0	282	2	US-08-795-303-19	Sequence 19, Appl
29	230	12.8	360	1	US-08-795-303-20	Sequence 20, Appl
30	230	12.8	360	2	US-08-674-612-4	Sequence 4, Appl
31	230	12.8	360	2	US-08-469-431-12	Sequence 12, Appl
32	230	12.8	360	2	US-08-250-975-12	Sequence 12, Appl
33	230	12.8	360	2	US-08-920-296-4	Sequence 4, Appl
34	230	12.8	360	2	US-08-605-002A-12	Sequence 12, Appl
35	230	12.8	360	2	US-08-950-449A-12	Sequence 12, Appl
36	230	12.8	270	2	PCR-7857-224B-31	Sequence 31, Appl
37	228	12.7	360	3	PCR-US94-10529-12	Sequence 12, Appl
38	228	12.7	360	2	US-08-674-612-3	Sequence 3, Appl
39	224	12.5	368	1	US-08-176-620A-15	Sequence 15, Appl

QY	124	IEEKDIKWLKSMSSGLEFIHSGOIIHRDIKPSNIFARDDITQPIIGDFIDCYDLKLPP	183
Db	106	LDPLVOKFTYQLVGVNFCFHSRRIIHRDLKPNQLIDKEGNLK--LADFGLARSGV--	162
QY	184	KDEPMAKII-DYGTGIYKAPELLIGITNYEYEDTWSIGIILITGLYSNFOSVLVKDK	242
Db	162	----PLRNYTHEIVTLWYRAPEVLLSGRSHYSTGVDIWSGCF-----AEMIRSP	208
QY	243	ELTNDSHVSDYLLNQIFENFGFNPNTDFEDELCEYNNENLHFKKENLQYPRKDWI	302
Db	209	LFPGDSIEIF--KIFQVLGTPNEEVGVTLLODYKSTFPWRWMDLHK-----	258
QY	303	ILPFCNDLMKEFTKMIYDRSKRITSKEILQ	335
Db	258	WPNGEEDAI-ELLSAMLVDPDPAHRSKRALQ	289

```

1  RESULT      4
2  PCT-US94-12913A-18
3  ; Sequence 18, Application PC/TUS9412913A
4  ; GENERAL INFORMATION:
5  ; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
6  ; APPLICANT: UNIVERSITY OF MASSACHUSETTS MEDICAL SCHOOL
7  ; APPLICANT: Karin, Michael
8  ; APPLICANT: Davis, Roger
9  ; APPLICANT: Bibi, Masahiko
10 ; APPLICANT: Lin, Anning
11 ; APPLICANT: Derlward, Benoit
12 ; TITLE OF INVENTION: ONCOPROTEIN PROTEIN KINASE
13 ; NUMBER OF SEQUENCES: 18
14 ; CORRESPONDENCE ADDRESS:
15 ; ADDRESSEE: Spensley Horn Jubas & Lubitz
16 ; STREET: 1880 Century Park East, Suite 500
17 ; CITY: Los Angeles
18 ; STATE: California
19 ; COUNTRY: USA
20 ; ZIP: 90067
21 ; COMPUTER READABLE FORM:
22 ; MEDIUM TYPE: Floppy disk
23 ; COMPUTER: IBM PC compatible
24 ; OPERATING SYSTEM: PC-DOS/MS-DOS
25 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
26 ; CURRENT APPLICATION DATA:
27 ; APPLICATION NUMBER: PCT/US94/12913A
28 ; FILING DATE:

```

Query Match 14.3%; Score 255.5; DB 3; Length 424;
Best Local Similarity 27.1%; Pred. No. 2.2e-15;
Matches 92; Conservative 49; Mismatches 112; Indels 87; Gaps 16;

[illegible]

```

RESULT      5
US-08-674-612-2
: Sequence 2, Application US/08674612
: Patent No. 5663313
: GENERAL INFORMATION:
: APPLICANT: Hawkins, Phillip R.
: APPLICANT: Au-Young, Janice
: APPLICANT: Guegler, Karl J.
: APPLICANT: Wilde, Craig G.
: TITLE OF INVENTION: A NOVEL HUMAN MAP KINASE HOMOLOG
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
: STREET: 3174 Porter Drive
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Fastseq Version 1.5
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/674,612
: FILING DATE: Herewith
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 60/000,722
: FILING DATE: 30-JUN-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J.
: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: PF-0036 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-855-0555
: TELEFAX: 415-845-4166
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 365 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: IMMEDIATE SOURCE:
: LIBRARY: Stomach
: CLONE: 214915
US-08-674-612-2

```

Query Match 14.0%; Score 251.5; DB 1; Length 365;
Best Local Similarity 26.4%; Pred. No. 4.1e-15;
Matches 91; Conservative 56; Mismatches 125; Indels 73


```

RESULT      9
US-08-874-347-18
; Sequence 18, Application US/08874347
; Patent No. 5863741
; GENERAL INFORMATION:
; APPLICANT: Limper, Andrew H.
; APPLICANT: Leof, Edward B.
; APPLICANT: Thomas, Charles F.
; APPLICANT: Gustafson, Michael P.
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS
; TITLE OF INVENTION: CARINII
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C., P.A.
; STREET: 60 South Sixth Street, Suite 3300
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/874,347
; FILING DATE: 13-JUN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ellinger, Mark S.
; REGISTRATION NUMBER: 34,812

```

```
; REFERENCE/DOCKET NUMBER: 07039/055001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-335-5070
; TELEFAX: 612-288-9696
; TELEX:
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 317 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-874-347-18

Query Match      13.7%; Score 246; DB 2; Length 317;
Best Local Similarity 25.8%; Pred. No. 1.le-14;
Matches 88; Conservative 69; Mismatches 122; Indels 62; Gaps 15;

QY    1 MKLSDDYYDKELIYNLSAISDIYTALD-KFNKLPVCLK---IVDEDFSLPPHSITHREIFL 56
       ::|||:: : : : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db    2 VELS DYQ-RQEKVGEGYGVVYKALDTKHNNRVVALKKIRLESEGGVFSTAI-REISLL 59
       |||:: : : : | : | : | : | : | : | : | : | : | : | : | : | : | :

QY    57 KTLKPHPNIIIEFYNDKIYDVILVT-KLYRYDLSQLIEITKYCKRTTTFYIINGNLWS 115
       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db    60 KEKKDD-----NIVRLYDIHSDSHKLIVFEFLDLCLKYMESIPQGV-GLGANM-- 110
       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

QY    116 NOYTLANEIEEKDKIKMLWKMSSGLETFHSGIIHRDIKPSNTFFARDDITOPIGDPTI 175
        ::|:: : : : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db    110 -----IKRFNMQLIRGIKCHSHRVLHRLDKPQNLLIDKEGNLK--LADFGL 154
        | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

QY    176 CYVDLKLPRKDEPPMAKYIDVSTGIYKAPELLIGTINYEYEDIWLSGILTLGYSENFS 235
        : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db    155 ARAFGVPLR-----AYTHEVVTLWYRAPEILLGGKQYSTGVDMSVGCFAECNR---- 206
        : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

QY    236 VLVKDDKEL-TNDSHVSVDLIILNQIFENFGTPNLTDFDELFCDEYNNEHLHFKEFL-Q 293
        | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db    206 -----KPLFGDSIDEIF--RIFRILGTPNIEIWPVDVNYLPDRFKSFPQWKPKLSE 256
        | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

QY    294 KYPRKDWDIILPRCNDLMMKEIFTMKIMRYDRSKRITSKEIL 334
        | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db    257 AVPSLDANGI-----DLLDQMLVDPDSRRISAKRAL 287
        | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

RESULT 10
US-08-874-347-24
; Sequence 247, Application US/08874347
; Patent No. 5863741
; GENERAL INFORMATION:
; APPLICANT: Limper, Andrew H.
; APPLICANT: Leof, Edward B.
; APPLICANT: Thomas, Charles F.
; APPLICANT: Gustafson, Michael P.
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS
; TITLE OF INVENTION: CARINI
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C., P.A.
; STREET: 60 South Sixth Street, Suite 3300
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/874,347
; FILING DATE: 13-JUN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
```

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ellinger, Mark S.
REGISTRATION NUMBER: 34,812
REFERENCE/DOCKET NUMBER: 07039/055001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-335-5070
TELEFAX: 612-288-9696
TELEX:
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 297 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-874-347-23

Query Match 13.3%; Score 238.5; DB 2; Length 297;
Best Local Similarity 25.1%; Pred. No. 4.6e-14;
Matches 84; Conservative 64; Mismatches 126; Indels 61; Gaps 14;
QY 7 YIDKELIYNSAISIYTAIDKFNLPVCLK---IVDEDFSLPPHSIHREIFILKTLKPHP 63
DB 4 YIKIEGEGTYGVVYKGRHRTTQGVAMKKIRLESEEGVPSTAI-REISLLKELR-HP 61
QY 64 NIEYFNDLKIYDD-VILVTKLYRYDLSOLIEITIKYCKRTTRFIYINGNLVSNQYTLAN 122
DB 62 NIVS-LQDVLQDSRLYLIEFELSMDLKKYLD-----SIPPGQF----- 100
QY 123 EIEKDKILKMSGSGLEFTHSGIHRDIKPSNIFPARDITQPIIGDFIDICYDLKLP 182
DB 100 -MSSSLVKSXYLIQLOGIVFCHSRRLVHRLKPNQLLI--DDKGTIKLADFGLARAFGI- 156
QY 183 PKDEPPMAKYD-DVSTGIYKAPELILGTNYEYEDIDWSLGIILTGLYSENFOSVLVKDD 241
DB 156 -----PIRVYTHVVTLYWRSPEVLLGSARYSTPVDIWSIGTIFAELATK----- 201
QY 242 KELTN-DSHVSDLYLLNQIFENFTGNLTDPELFCDEYNENLHFKFNLOKYPKDW 300
DB 201 KPLFHGDSIEDQLF---RIFRALGTPNNEVWPEVESLDQYKNTFPKWKPGSLASHVKN-- 256
QY 301 DIILPRCNDLMLKEIFTKMIRYDRSKRITSKEILQ 335
DB 256 -----LDENGLDLSKMLVYDPKAKRISGMALK 283

RESULT 12
US-08-176-620A-16
Sequence 16, Application US/08176620A
Patent No. 5595904
GENERAL INFORMATION:
APPLICANT: Boulton, Teri G.
APPLICANT: Cobb, Melanie H.
APPLICANT: Yancopoulos, George D.
APPLICANT: Nye, Steven
APPLICANT: Panayotatos, Nikos
TITLE OF INVENTION: A Family of Map2 Protein Kinases
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/176,620A
FILING DATE: 03-JAN-1994
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 6526-123
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 297 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-176-620A-16

Query Match 13.3%; Score 237.5; DB 1; Length 297;
Best Local Similarity 25.4%; Pred. No. 5.6e-14;
Matches 85; Conservative 60; Mismatches 128; Indels 61; Gaps 14;
QY 7 YIDKELIYNSAISIYTAIDKFNLPVCLK---IVDEDFSLPPHSIHREIFILKTLKPHP 63
DB 4 YIKIEGEGTYGVVYKGRHRTTQGVAMKKIRLESEEGVPSTAI-REISLLKELR-HP 61
QY 64 NIEYFNDLKIYDD-VILVTKLYRYDLSOLIEITIKYCKRTTRFIYINGNLVSNQYTLAN 122
DB 62 NIVS-LQDVLQDSRLYLIEFELSMDLKKYLD-----SIPPGQY----- 100
QY 123 EIEKDKILKMSGSGLEFTHSGIHRDIKPSNIFPARDITQPIIGDFIDICYDLKLP 182
DB 100 -MSSSLVKSXYLIQLOGIVFCHSRRLVHRLKPNQLLI--DDKGTIKLADFGLARAFGI- 156
QY 183 PKDEPPMAKYD-DVSTGIYKAPELILGTNYEYEDIDWSLGIILTGLYSENFOSVLVKDD 241
DB 156 -----PIRVYTHVVTLYWRSPEVLLGSARYSTPVDIWSIGTIFAELATK----- 201
QY 242 KELTN-DSHVSDLYLLNQIFENFTGNLTDPELFCDEYNENLHFKFNLOKYPKDW 300
DB 201 KPLFHGDSIEDQLF---RIFRALGTPNNEVWPEVESLDQYKNTFPKWKPGSLASHVKN-- 256
QY 301 DIILPRCNDLMLKEIFTKMIRYDRSKRITSKEIL 334
DB 256 -----LDENGLDLSKMLIYDPKAKRISGMAL 282

RESULT 13
US-08-461-985-16
Sequence 16, Application US/08461985
Patent No. 5872006
GENERAL INFORMATION:
APPLICANT: Boulton, Teri G.
APPLICANT: Cobb, Melanie H.
APPLICANT: Yancopoulos, George D.
APPLICANT: Nye, Steven
APPLICANT: Panayotatos, Nikos
TITLE OF INVENTION: A Family of Map2 Protein Kinases
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

```

; SOFTWARE: PatentIn Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,985
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 800
;
; PRIOR APPLICATION NUMBER: US 08/176,620
; FILING DATE: 03-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistock, S. Leslie
; REGISTRATION NUMBER: 19,872
; REFERENCE/DOCKET NUMBER: 6526-123
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 297 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
;
; US-08-461-985-16
;
; Query Match 13.3%; Score 237.5; DB 2; Length 297;
; Best Local Similarity 25.4%; Pred. No. 5.6e-14;
; Matches 85; Conservative 60; Mismatches 128; Indels 61; Gaps 14;
;
; QY 7 YDKELIYNSAISDIYTAIDKFNLPVCLK--IVDEDFSLPPHSIHRIFILKTLKPHP 63
; | | | | | | | | | | | | | | | | | | | | | | | | | | | |
; Db 4 YTKIEKIGEGTVGVYKGRHKTGTGVAMKIRLESEEGVPSTAI-REISLLKELR-HP 61
; | | | | | | | | | | | | | | | | | | | | | | | | | | | |
; QY 64 NIEVFNDLKIYDD-VILVTKLYRYDSLQLEIYTKYKRTTRFYINGNLVSNQYTLAN 122
; || | | | | | | | | | | | | | | | | | | | | | | | | | | |
; Db 62 NIVS-LQDVLMODSRLYLIFEFLSMDLKKYLD-----SIPPGQY----- 100
; | | | | | | | | | | | | | | | | | | | | | | | | | | | |
; QY 123 EIEEKDKIWLKSMSSGLEFTHSQGIHHRDKPSNIFFARDITQPIGDFDICYDKLP 182
; : : : : : : : : : : : : : : : : : : : : : : : : : : : :
; Db 100 -MDSSSLYKLYQILQGIQVFCFHSRVLHRDLKPNQLLI--DDGKTGLADFGIARAFGI- 156
; | | | | | | | | | | | | | | | | | | | | | | | | | | | |
; QY 183 PKDEPPMAXIYD-VSTGIYKAPELIIGITNTYEYEDISLGIILTGLYSNFOSVLKDD 241
; | | | | | | | | | | | | | | | | | | | | | | | | | | | |
; Db 156 ----PIRYTHVTVLWYRSEVLGSRYSTPVDIWSIGTIFAEALATK----- 201
; | | | | | | | | | | | | | | | | | | | | | | | | | | | |
; QY 242 KELTN-DSHVSDLYLLNQIFENFGTNPNTDEDELFCDENNNENLHKKNLQYPRKDW 300
; | | | | | | | | | | | | | | | | | | | | | | | | | | | |
; Db 201 KPLFGHDSQIDQLF---RIFRALGTPNNEVMPVESLQDYKNTFPKWKPGSLASHVKN--- 256
; | | | | | | | | | | | | | | | | | | | | | | | | | | | |
; QY 301 DILPRCNDLAKIEFTKMIRYDRSKRITSKEIL 334
; | | | | | | | | | | | | | | | | | | | | | | | | | | | |
; Db 256 -----LDENGLDLSKRLIYDPAKRISGKMAL 282
; | | | | | | | | | | | | | | | | | | | | | | | | | | | |
;
; RESULT 14
; US-08-176-620A-14
; Sequence 14, Application US/08176620A
; Patent No. 5595904
; GENERAL INFORMATION:
; APPLICANT: Boulton, Teri G.
; APPLICANT: Cobb, Melanie H.
; APPLICANT: Yancopoulos, George D.
; APPLICANT: Nye, Steven
; APPLICANT: Panayotatos, Nikos
; TITLE OF INVENTION: A Family of Map2 Protein Kinases
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
;

```

STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,985
FILING DATE: 05-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/176,620
FILING DATE: 03-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 6526-123
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 353 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-461-985-14

Query Match 13.2%; Score 236.5; DB 2; Length 353;
Best Local Similarity 25.7%; Pred. No. 8.7e-14;
Matches 90; Conservative 62; Mismatches 125; Indels 73; Gaps 16;

QY	4	SDYYIDKELYNLSAISDIYTAIDKFNLPVCLKIVDEFSLPSPHSIH--REIFILKTLKP	61
Db	11	SDFQ-KSLGEGAYGVCSATHKPTGEIVAIKKI-EPEDKPLFALRLREIKLKHKF-	68
QY	62	HPNIIIEYFN-----DLKIYDDVILTKLYRDLSQLIEITKYCKRTTRIYINGNLVSN	116
Db	68	HENIIIFNIQRDPSFNFNVEVYIQLMOTDLHRVIS-----	106
QY	117	QYTLANEIEKDKILMLKSMSSGLEFIHQSGIIHRDIKPSNIFF-ARDDITQPIIGDFDI	175
Db	106	-----TQMLSDDHQYFYIQTIRAVKVLGHSNVIRDLKPSNLLINSNCDLK--VCDFGL	158
QY	176	-----CYDLKLPKDEPPMAKYIDVSTGIYKAPELIIGITNVEYEDISLGIITGL	228
Db	159	ARIIDESAADNSEPTGOQSGMTEY--VAIRWYRAPEVMTLSAKYRAMDVWSCGCI LAEL	216
QY	229	YSENFQSVLVKDDKELTNDSHVSDLYLLNQIFENFGTPNLTDPELFCDEYNENLHEK	288
Db	217	FLR--RPIFGRDYR-----HQLLLIFGICTPH---SDNDLCIESPRAREYIK	261
QY	289	KFNQYKYPKRDWDIILPROND---DLMKEIFTKMIRYDRSKRITSKEILQ	335
Db	262	--SLPMYPAAPLEKMFPRVNPKNIGIDLLQ----RMLVDFPAKRITAKEALE	305

Search completed: November 6, 1999, 08:13:56
Job time: 1504 sec

Insulin-stimulated	44	49.2	4.8	2260	1	T27730
Mutant insulin-sti	45	49.2	4.8	2260	1	T27731

Result No.	Score	Query Match	Length	DB	ID	Description
1	1007.8	98.9	1020	1	X27495	C.albicans Cdk act
2	71.2	7.0	1671	1	X24134	50 kD subunit of S
3	66.2	6.5	4590	1	N60472	Sequence encoding
4	64.2	6.3	1671	1	X24134	50 kD subunit of S
5	60	5.9	110000	1	V21209_03	Continuation (4 of
6	59.6	5.8	1864	1	N71405	Sequence of ANS-1
7	58.6	5.8	605	1	T31530	Human 3' apolipop
8	58.2	5.7	19124	1	T72882	Plasmidium var-7 g
9	57.2	5.6	605	1	T31530	Human 3' apolipop
10	56.8	5.6	5760	1	N50530	Sequence encoding
11	56.6	5.6	5563	1	X08941	GARA-1 promoter re
12	56.2	5.5	8920	1	Q62924	Carbamoyl-phosphat
13	54	5.2	2503	1	Q53480	PNPX30 xylanase CD
14	53.4	5.2	19124	1	T72882	Plasmidium var-7 g
15	53.2	5.2	9789	1	T41852	cDNA encoding Plas
16	53	5.2	1070	1	T64450	MOC1 protein. New
17	53	5.2	2483	1	T63319	Yeast mitochondria
18	52.6	5.2	110000	1	V21209_15	Continuation (16 o
19	51.8	5.1	51952	1	X26084	Tomato pest resist
20	51.4	5.0	5203	1	V62901	Yeast mitochondria
21	51.4	5.0	110000	1	X20248_04	Continuation (5 of
22	51.4	5.0	18277	1	X20249	Borrelia burgdorfe
23	51.2	5.0	110000	1	X20248_05	Continuation (6 of
24	51	5.0	110000	1	V21209_13	Continuation (14 o
25	51	5.0	110000	1	V21209_14	Continuation (15 o
26	50.8	5.0	1230	1	N90223	Malaria-specific D
27	50.8	5.0	4590	1	N60472	Sequence encoding
28	50.4	4.9	6124	1	Q35568	Sequence encoding
29	50.4	4.9	8820	1	Q62924	Carbamoyl-phosphat
30	50.4	4.9	3974	1	T43360	Cotton FbL2at2-G2A
31	50.2	4.9	10715	1	X20248_09	Continuation (10 o
32	50	4.9	1644	1	V26286	Recombinant botuli
33	50	4.9	3509	1	V26289	Recombinant botuli
34	50	4.9	2574	1	V26290	Recombinant botuli
35	50	4.9	2574	1	V26292	Recombinant botuli
36	50	4.9	1711	1	T33136	Plasmidium berghel
37	50	4.9	5852	1	V30579	Clostridium botuli
38	49.8	4.9	3876	1	Q11710	Dictyostelium plas
39	49.8	4.9	110000	1	X20248_06	Continuation (7 of
40	49.8	4.9	111309	1	X30250	Borrelia burgdorfe
41	49.6	4.9	6124	1	Q35568	Sequence encoding
42	49.6	4.9	1711	1	T33136	Plasmidium berghel
43	49.4	4.8	1864	1	Q78892	Aspergillus nidula

437

back date all prior

Db 625 TTATCTATTAAATAAATAATATTAGATAAATTAAGTTGAATATTAGATAGTTTATAATTT 566
QY 672 TTTGACTGGTTTATATATTCAGAAATTTTCAAGAGTTTGTAGTCAAGAGATGATAAAGAAAT 731
Db 565 AAATTTATTTATATATTTTAAATATTCATTTATTTTAAAGAAATTTATTTATTTATTA 506
QY 732 GACTAATGATCTCATGTTAGTGTATATATTTTAAATCAATATTTTGAATATTTTGG 791
Db 505 TTATTAATATCTATTTCTATTTTATTTTAAATATAAATAAAGATATAAT 446
QY 792 TACACCCCAATTAACGATTTTGAAGATGAATTTATTTGTGATGA 836
Db 445 AAATTAATCTATAATATTTTGAATTTTATTAATCTTTTATTA 401

RESULT 5
V21209_03/c
Continuation (4 of 17) of V21209 from base 300001 (Methanococcus jannaschii circular chr
WP Sequence split into 17 fragments LOCUS V21209 Accession V21209
Fragment Name Begin End
WP V21209_00 1 110000
WP V21209_01 100001 210000
WP V21209_02 200001 310000
WP V21209_03 300001 410000
WP V21209_04 400001 510000
WP V21209_05 500001 610000
WP V21209_06 600001 710000
WP V21209_07 700001 810000
WP V21209_08 800001 910000
WP V21209_09 900001 1010000
WP V21209_10 1000001 1110000
WP V21209_11 1100001 1210000
WP V21209_12 1200001 1310000
WP V21209_13 1300001 1410000
WP V21209_14 1400001 1510000
WP V21209_15 1500001 1610000
WP V21209_16 1600001 1664976

Query Match 5.98; Score 60; DB 1; Length 110000;
Best Local Similarity 44.98; Pred. No. 0.023;
Matches 317; Conservative 0; Mismatches 380; Indels 9; Gaps 2;

QY 261 TGATTTGAGTCAATGATTTGAAATTAACAAATATGTTAAACGACACACAGATTTATTTA 320
Db 85823 TGATATTTGAATATTTATGAATGATTAAGGATAAAGAACTAAAGAGGTTATATGA 85764
QY 321 TGGTATTAATCGTAACTCTGTTAGTAAATCAATATACACTTGCTAAATGAAGAAA 380
Db 85763 AAAAAGATGATGAAATTTGTTGTAATATAAATATAAAGTTGATATTTATATGTTGA 85704
QY 381 AGATATCAATATATGTTTAAATCAATCAGTTCCAGGACTTGAATTTATTCATTCACAAG 440
Db 85703 ATTTAATCGTATTTAGATTTTGTAAATATGGAATATCAATTCGTATCCGTTT 85644
QY 441 GATATTCATCGTGATATAAACCAGTAATATTTTCITGCGCGGATGATATACACA 500
Db 85643 TAATGGAATGCTATTTTAGAGTCATAACTTTTAAAGGAAGTGTTTTATTTATAAGAAC 85584
QY 501 ACCGATTTATGGAGTTTGTATATTTGTTA-----TGATTTAAACATGCCACCTAAAGA 554
Db 85583 AATTTTATGGAGATGTTGATTTTATAGACACAATTTTGAAGAAAATGTTTATTTAG 85524
QY 555 TGAACCCCTATGGGAAATATATGATGTATCTACAGGTATTTTAAAGCACACAGAAAT 614
Db 85523 TGTACCCCTTTTAAAGGAATATTTAATTTTAGTGGCAATTTTAAATAAGAAATC 85464
QY 615 GATTCCTGGTAACTAATTAATGAATTAATTAATTTGATTTTGGTCATTTGGGTATAATTT 674
Db 85463 TCAITTTTAAAGTACAACTTTTGAAGGAATTAATTTTATTTAGTGTCAACACTTTTAAAT 85404
QY 675 GACTGGTTTATATTCAGAAAATTTTCAAGTGTTTTAGTCAAGATGATAAAGAAATTCAC 734

Db 85403 AGCAGAAATTTTAAATCAACATTTAAATCTCATGTATATTTTGATGATATATCATTTAA 85344
QY 735 TAATGATCTCATGTTAGTGTATTTATTTAAATCAATATTTTGAATATTTTCGGTAC 794
Db 85343 TTTGTTCTCTTACTGATTTAGATTTAGAGATGATGATCATTTTAAAAAATAGATAA 85284
QY 795 ACCCAATTTAACTGATTTTGAAGATGAATTTTGTGATGAATATTAATCAAACTT 854
Db 85283 AGAAATTTTAAAGTTTAGCAAT---ATTCTTAAAACTCAATTTCTTAAACAAACACAC 85227
QY 855 GCATTTTAAAAAATTTCAATTTACAAAATATCTAGAAAAGATTTGGGATATTTTATACC 914
Db 85226 AACGATAGAAAATTTCCAAATATCAAAAACATCATTTTAAAAACAGATGTTAGAGAGT 85167
QY 915 TCGATGCAATGATGATTTAATGAAGAAAATTTTACCAAGATGATT 960
Db 85166 GCTATTATGATGCTTTAAAAAAGAGAAATTTTAAAGTCATAAAATTT 85121

RESULT 6
N71405/c
ID N71405 standard; DNA; 1864 BP.
AC N71405;
DT 18-APR-1991 (first entry)
DE Sequence of ANS-1 which increases transformation efficiency.
KW Enzyme; fungal expression vector; Aspergillus expression vector;
KW Trichoderma; ds.
OS Mucor miehei.
PN EP-215594-A.
PD 25-MAR-1987.
PE 27-AUG-1986; 306624.
PR 29-AUG-1985; US-771374.
PR 07-JUL-1986; US-882224.
PA (GENE-) GENENCOR INC.
PI Cullen D, Gray GL, Hayenga KJ, Lawlis VB;
DR WPI: 87-095049/14.
PT New DNA sequences for expressing polypeptide in filamentous fungi
PT - with secretion of prod. from the cells, and new vectors and
PT transformed fungi
PS Example; Fig 13; 45pp; English.
CC A DNA sequence coding for a heterologous polypeptide which can be
CC expressed in and secreted from filamentous fungi is claimed. Pref.
CC the DNA sequence codes for bovine preprochymosin, M. meihei
CC preprocarboxyl protease or A. niger preproglucoamylase. Also new
CC are vectors consisting of the DNA sequence plus an operably-linked
CC signal sequence. The vectors may also include a sequence which
CC increases transformation efficiency, e.g. ANS-1.
CC Sequence 1864 BP; 786 A; 210 C; 44 G; 732 T;
SQ

Query Match 5.88; Score 59.6; DB 1; Length 1864;
Best Local Similarity 42.48; Pred. No. 0.03; 532; Indels 8; Gaps 3;
Matches 397; Conservative 0; Mismatches 532; Indels 8; Gaps 3;

QY 86 ATAACTTACCAGTATCTCTTAAATAGTTGATGAAGATTTCCAGTCTCCACCACTCAA 145
Db 1143 ATATCTGNTAATATATTTTANTATTTTAAATATCTAATATAGTAAAGTATAGTAGG 1084
QY 146 TCCATCGGAAATTTTATACCTTAAACTTTGAAACCACTCCCAACATATTTGAATTT 205
Db 1083 TTTTATAATTTATAATAATAATAATAATTTTAAAGTTATAAGNTNTAATTTTAAAT 1024
QY 206 TTAATGAATTTAAATTTATGATGATGTTATATAGTACCACCAATTTGATGTTATGAT 265
Db 1023 TTTAGCTATTAAATTAATTAATTAATTTTATTAATAATTTTATAATTTATTTTCGAG 964
QY 266 TGAGTCAATTTGATTTGAAATTTACAAAATATTTGTAACGACCAACACAGATTTATTTATGGA 325
Db 963 GGAGAGTTGTTTAAATAATAATAATAATAATAATAGTAGTTTATTTATATATATATAATA 904
QY 326 TTAATGGTAACTCTGTTAGTAAATCAATATACACTTGCTAAATGAATTTGAAGAAAAGATA 385
Db 903 ATATTTTNTAGTAGTGTATTTAATAAAGAGCTTTAAAAATTTATTTCTTTTAGGAT 844

PT New malaria vaccines - contains cysteine-rich DBL family protein
 PT binding domains homologous domains of the Duffy and sialic acid
 PT binding proteins

PS Claim 4: Page 56-61; 96pp; English.

CC This sequence represents the var-7 gene of Plasmodium. Var-7 belongs to
 CC the Duffy binding like (DBL) family of genes which have homology to the
 CC Duffy antigen binding protein (DABP) and sialic acid binding protein
 CC (SABP) conserved regions (see T72889 and T72888 respectively). The var
 CC family of genes modulate cytoadherence and antigenic variation of
 CC Plasmodium infected erythrocytes. SABP and the Duffy antigen binding
 CC protein (DABP) are soluble proteins that appear in the culture
 CC supernatant after infected erythrocytes release merozoites. DABP and SABP
 CC mediate the binding of merozoites and schizonts to the erythrocyte
 CC surface. These proteins are necessary for erythrocyte invasion by the
 CC parasite. This sequence can be used in the compositions of the invention.
 CC The compositions are for the treatment and prevention of malaria, and
 CC comprise either a nucleotide sequence or encoded polypeptide of the
 CC var-1, var-2, var-3 or var-7 genes of the DBL gene family, a family of
 CC genes having homology with conserved regions of DABP and SABP. The
 CC compositions are used for the treatment and prevention of malaria. They
 CC are also used in the preparation of vaccines for inducing a protective
 CC immune response in a mammal to Plasmodium merozoites (especially
 CC Plasmodium falciparum or Plasmodium vivax).
 CC Sequence 19124 BP; 7824 A; 2190 C; 2790 G; 6320 T;

Query Match 5.7%; Score 58.2; DB 1; Length 19124;

Best Local Similarity 45.2%; Pred. No. 0.047; Mismatches 543; Indels 20; Gaps 6;

QY 5 AGTTGTCAGATTATTATATAGACAAGAAATTAATTTACAAATAGTCCCTTTCTGATATAT 64
 DB 7195 AGTAGTATATATATGTTAGATAGATAGATAGATAGATAGATAGATAGATAGAT 7136
 QY 65 ATACGGCTATGTAAGTTTAATTAACCTTACCAGTAGTCTTAAATAGTTGATGAGAT 124
 DB 7135 GTAAGGTATAATTTATGTTATTAACAATATATATAATTTGTTATATATATATATA 7076
 QY 125 TCAGTCTTCACACACATTCATCCAGAGAAATTTTATCTTAA----AACTTTGAA 179
 DB 7075 TTAATAGTCTGCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7016
 QY 180 ACCACATCCAAACATATGTAATTTATTTAATGATCTTAAATTTAATGATGATGAT 239
 DB 7015 TATGATATATAATAATTAATTTATGTTTATATATTTTAAAAAATATATATATAT 6936
 QY 240 AGTCACCAATTTGATCGTTATGATTTGATGATGATGATGATGATGATGATGATG 299
 DB 6955 AATAAATTTAATAAATTTAATAATTTCTAACAAAAAATAATATATCAGAAATATAT 6896
 QY 300 ACGAACACACGATTTATTTATGTTATGTTATGTTATGTTATGTTATGTTATGTTAT 359
 DB 6895 ATTTTATGATCTTTATTTATCTATTTAATTAATATATATATATATATTTTATGTT 6836
 QY 360 TGCTAATGAATTTGAAGAAAGATATCAATTTATGTTGTTAAATCAATGATGATGAT 419
 DB 6835 ATTTTAAAGTAAATTTAATAATGAGAAAAAATAATGAGAAATACAAACATATAAAA 6776
 QY 420 TGAATTTATTCATTCACAGGGATATTCATCGTATATATAAACCCAGTAATTTCTT 479
 DB 6775 AGTATATATGCAAGCGTTTATATATTTAATTAATTAATTAATTAATTAATTAAT 6716
 QY 480 TGCCCGGATGATATACACACCGATTTATTTGGAGATTTTGTATTTGTTATGATTTAA 539
 DB 6715 TGACTTTATTTATTTAT 6656
 QY 540 ACTGCCACCTAAAGATGAACCCCTATGCGGATATATTTGATGATCTATCAGAGTATTT 599
 DB 6655 TGTTCTTCGTTCTCTCTATCTA-TATTTATCTATCATATATATATATATATATAT 6597
 QY 600 TAAACACACCAAGATTTGATCTTGTGTAACCTAATTTATGAAATGAAATGATTTGGTC 659
 DB 6596 ATTGATATAGATACATATCTTCTGTTGTTATATATATATATATATATATATAT 6541

QY 660 ATTGGGTATATATTTGACTGGTTTATATATTCAGAAAAATTTTCAAGTGGTTTATAGTCAAAGA 719
 DB 6541 ATTATTTAATTTTTCGTTTAT 6482
 QY 720 TGAATAAGAAATGACTAATGATCTCTCAATGTTAGTATTTATTTTAAATCAATATTT 779
 DB 6481 CAAAAAATAATGATAAATAAATAACAGAGAAATAATATATATATATATATATATATATAT 6422
 QY 780 -----TGAAAAATTTTCGTACACCAATTTAACTGATTTTGAAGATGAATATTT---TGT 831
 DB 6421 ATTAATAAATAAGTTTATATCAATTTGTTTGTGTTATTTTATGTTATTCATGCAATTT 6362
 QY 832 GATGAATATATAATATGAAACTTTCATTTTAAATAATCAATTTTCAAAAATATCCCTAGA 891
 DB 6361 TATGAATTTCAAAATTTTATTTATATATAAATAAATAAGTAAAAATACATTTATATAA 6302
 QY 892 AAAGATTGGGATATATTTTACCTCGATGCAATGATGATTTAATGAAGAAATTTTATACC 951
 DB 6301 TATATATATATCAATATAGATTTATTAATAAATGTTCAATGTTCTATATATATATATAAT 6242
 QY 952 AAGATGATTAGATATGCAAGTAAAGAAATAACTTCTAAAGAAATCTTACAAATTAATG 1011
 DB 6241 GAAATATTTGTTATCT 6183
 QY 1012 TTGGATT 1018
 DB 6182 TTAAGT 6176

RESULT 9

ID T31530 standard; cDNA; 605 BP.
 AC T31530;
 DT 15-SEP-1996 (first entry)
 DE Human 3' apolipoprotein B SAR element clone Rh32.
 KW Erythropoietin; EPO; anaemia; gene therapy; vector;
 KW scaffold attachment region; SAR element; apolipoprotein B;
 KW transgenic animal; ss.
 OS Homo sapiens.
 FN WO9619573-A1.
 PD 27-JUN-1996.
 PF 18-DEC-1995; CA0696.
 PR 19-DEC-1994; US-358918.
 PA (CANG-) CANGENE CORP.
 PI Delcove G;
 DR WPI; 96-309587/31.
 PT Recombinant DNA molecule expressing mammalian erythropoietin -
 PT useful to transform cell lines, and for gene therapy, e.g. of
 PT anaemia and other red blood cell disorders
 PS Claim 7; Page 59-60; 84pp; English.
 CC Human apolipoprotein B (apoB) scaffold attachment region (SAR)
 CC element clones Rh32 (T31530) and Rh10 (T31531) respectively carry
 CC the 3' human apoB SAR element and the distal 1212 bp 5' human apoB
 CC SAR element and 1317 bp proximal sequence. These SAR elements
 CC co-map with the boundaries of the human apoB gene chromatin
 CC domain. A novel recombinant DNA molecule adapted for transfection
 CC of a host cell comprises an erythropoietin (EPO) cDNA (T31529) or
 CC genomic clone (T31532) operably linked to an expression control
 CC sequence and to the 5' and 3' SAR elements. The SAR elements
 CC increase expression of the recombinant EPO in stable, long-term
 CC mammalian cell cultures.
 SQ Sequence 605 BP; 278 A; 18 C; 35 G; 274 T;

Query Match 5.6%; Score 57.2; DB 1; Length 605;

Best Local Similarity 46.3%; Pred. No. 0.077; Mismatches 188; Conservative 0; Mismatches 218; Indels 0; Gaps 0;

QY 571 AATATATGATGATCTACAGGTATTATATAAGACACAGCAATTCATCTTGGTAACT 630
 DB 63 AATATATATATAAATAATTTTAAATAATTTATATATATATATATATATATATATAT 122

QY	631	AATTAGTAATGAAGAATTGATATTTGGTCATTTGGGTATAAATTTTGACTGGTTTATATATCA	699
Db	123	AAATATTTATATATTTAAATATATTTTATAATTTAAATATTTTATAATTTAAATATTTTATAATTA	182
QY	691	GAAATTTTCAAAGTGTTTGTAGTCAAGAGTGATAAAGAATTGACTAATGATCTCAATGTT	750
Db	183	AAATATTTATAAATTTAAATATTTTATAATTTAAATATTTTAAATATTTTATAATTA	242
QY	751	AGTGATTTATATTTATTAAATCAAAATATTTGAAAAATTTCCGTACACCCAAATTTAACTGAT	810
Db	243	AAATATTTATATTTAAATATATTTTATAATTTAAATGTTTATAATTTACATATTTTATAATTA	302
QY	811	TTTGAAGATGAATTTATTTTGTGATGAATATAAATAATGAAACCTGTCATTTTAAAAAATTC	870
Db	303	AAATGTTTATAATTTACATATTTTATAATTTAAATGTTTATAATTTACATATTTTATAATTA	362
QY	871	AATTACAAAATAATCTGAGAAAGATTTGGGATATATTATTTCCTCGATCGCAATGATGAT	930
Db	363	AAATGTTTATAATTTACATATTTTATAATTTAAATTTACATATTTTAAAGTATTTATAATTACATA	422
QY	931	TTAATGAAGAATAATTTTACCAGATCATGATATGATGATCGAAGTA	976
Db	423	TTTTTAAATTAAGTATTTTATAATTTACATATTTTATAATTTAAAGTA	468
RESULT 10			
N50530/c			
ID	N50530 standard; DNA; 5760 BP.		
AC	305030;		
DT	30-SEP-1991 (first entry)		
DE	Sequence encoding the P195 protein of Plasmodium falciparum.		
KW	Malaria vaccine; epitope; antigen; immunogen; ss.		
OS	Plasmodium falciparum.		
FH	Key	Location/Qualifiers	
FT	Cds	216..5179	
FT		/*tag= a	
PN	EP-154454-A.		
PD	11-SEP-1985.		
PF	21-FEB-1985; 3011173.		
PR	22-FEB-1984; GB-004692.		
PR	26-SEP-1984; GB-024340.		
PR	21-FEB-1985; GB-004429.		
PI	(WELL) WELLCOME FOUNDATION LTD.		
PI	Holder A, Sandhu J, Odink K, Lockyer M, Riveros-Moreno V;		
DR	WFI; 85-224845/37.		
DR	P-PSDB; P50777.		
PT	Cloned DNA sequence encoding plasmodium falciparum protein -		
PT	useful for expressing the protein for use in vaccines against		
PT	malaria		
PS	Claim 2; Fig 1; 51pp; English.		
CC	The sequence encoding the P195 protein of Plasmodium falciparum		
CC	(N50530) and a peptide comprising at least one of its epitopes		
CC	(see P50777) are claimed. Also claimed is a vaccine for inducing		
CC	immunity to malaria comprising the novel peptide or p195 or a		
CC	peptide comprising at least one epitope when derived from the new		
CC	DNA sequence, together with a carrier.		
SQ	Sequence	5760 BP; 2565 A; 630 C; 725 G; 1840 T;	
Query Match			
Best Local Similarity 5.6%; Score 56.8; DB 1; Length 5760;			
Matches 276; Conservative 0; Mismatches 337; Indels 7; Gaps			
QY	237	ATTAGTCACCAAAATGTCATCGTTTATGATTTAGTCAATGATGAAATACAAAATATG	296
Db	5758	ATAATCTCTTTAGTT	5699
QY	297	TAAACGAACAACACCGATTTATTATTGGTATTAAATGTTAGTAACTCTGTAGTAATCAATATAC	356
Db	5698	TAAATGTTATTATTAAAGTTATTTTATAATGTTGAATGTTTATAAAAAAATAAAATTT	5699
QY	357	ACTTGCTAATGAATGTGAGAAAGAAATATCAAAATTTAGTTGTTTAAAAATCAATGCTCAGG	416

atches 119; Conservative 2; Mismatches 106; Indels 0; Gaps 0;	
QY	625 ATAACTAATTATGAATATGAATATTTGGTCATATTTGGGTATATTTGACTGGTTTA 684
DB	2194 ATTTTAAATATGAAGTCGAGTGTTCATGTATTTATTTGAGCATACTGTACCAAGCACCA 2135
QY	685 TATTCAGAAAATTTCAAAGTGTTCATGTCATGATGATGATGATGATGATGATGATGATGAT 744
DB	2134 TGTTATTGATTACATGAGGTGTTCATGATGATGATGATGATGATGATGATGATGATGATGAT 2075
QY	745 CATGTTAGTGAATTTATTTATTTAAATCAATATTTGAAATTTTCGGTACACCAATTTA 804
DB	2074 TAAATTAAGAACTAAACITTCGATATATAATACTCAATATATCTGGATAAGATAATACT 2015
QY	805 ACTGATTTTGAAGATGAATATTTTGGTATGATGATGATGATGATGATGATGATGATGATGAT 851
DB	2014 ACTTATTTGTATATTAATTTTGGTCCAAATATAATGTTGCATA 1968
RESULT 12	
ID	Q62924 standard; cDNA; 8920 BP.
AC	Q62924;
DT	06-DEC-1994 (first entry)
DE	Carbamoyl-phosphate-synthetase II.
KW	Carbamoyl-phosphate-synthetase II; CPSII; pscpsII gene;
KW	malaria; ss.
OS	Plasmodium falciparum.
FH	Key
FT	cds
FT	Location/Qualifiers
FT	1226..8401
FT	/*tag= a
FT	/EC_number= 6.3.5.5
PN	W09412643-A.
PD	09-JUN-1994.
PF	02-DEC-1993; AU0617.
PR	03-DEC-1992; AU-006206.
PR	16-DEC-1992; AU-006380.
PA	(UNIX) UNISEARCH LTD.
PI	Flores MV, Osullivan WJ, Stewart TS;
DR	WPI; 94-200271/24.
DR	P-PSDB; R55694.
PT	Nucleic acid encoding carbamoyl phosphate synthetase II -
PT	isolated from Plasmodium falciparum, used to develop prods. for
PT	the treatment of malaria.
PS	Disclosure; Page 6-16: 31pp; English.
CC	The cDNA sequence encoding the carbamoyl-phosphate-transferase II
CC	(CPSII) of Plasmodium falciparum was determined. The cDNA encodes
CC	a protein that includes 2 insert sequences not found in other CPSII
CC	proteins. The first separates the putative structural subdomain and
CC	the glutaminase subdomain of the glutamine-amidotransferase subunit
CC	of CPSII, while the second separates 2 ATP binding subdomains of the
CC	CPSII subunit, CPSa and CPSb.
SQ	Sequence 8920 BP; 3836 A; 774 C; 1232 G; 3078 T;
Query Match 5.5%; Score 56.2; DB 1; Length 8920;	
Best Local Similarity 45.4%; Pred. No. 0.1;	
Matches 328; Conservative 0; Mismatches 383; Indels 12; Gaps 3;	
QY	153 AGAAATTTTATACCTAAATTTGAAACCAACATCCACATATTTGAATATTTAATGA 212
DB	5854 AGGGAATAATAAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5913
QY	213 TCCTAAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 272
DB	5914 CCCTGGTAATCTTATATGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 5973
QY	273 ATTGATGAAATACAAAATATTTAAACGCAACACAGATTTATTTGTTAATGCG 332
DB	5974 GATGAAGAGTAAATAGATAACGAAAATTTAAATGATGATGATGATGATGATGATGATGAT 6033
QY	333 TAACTTTGTTAGTAAATACATACACTTCCTAATGAATTG-----AGAAAAAGAT 384

. Sat Nov 6 11:07:18 1999

us-09-072-994-13.rng

Page 11

Search completed: November 6, 1999, 07:50:24
Job time: 5127 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 6, 1999, 08:22:25 ; Search time 12.23 seconds
(without alignments)
1110.564 Million cell updates/sec

Title: US-09-072-994-14
Perfect score: 1791
Sequence: 1 MKLSDYYDKELIYNSAID.....IRYRSKRITSKEILQLMLD 339
Scoring table: BLOSUM62
Searched: 122810 seqs, 40065486 residues
Database : PIR_60.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	276.5	15.4	349	2	S68675	mitogen-activated
2	273	15.2	296	2	S24386	protein kinase (EC
3	267	14.9	368	2	S56225	hypothetical prote
4	265	14.8	308	2	S53538	protein kinase (EC
5	264	14.7	297	2	A37871	protein kinase (EC
6	262	14.6	391	2	A30319	casein kinase II (
7	262	14.6	391	2	S21335	casein kinase II (
8	262	14.6	384	2	B30319	casein kinase II (
9	262	14.6	391	2	A38611	casein kinase II (
10	262	14.6	298	2	A41227	protein kinase (EC
11	261	14.6	298	2	A44878	protein kinase (EC
12	261	14.6	306	2	S39151	csk1 protein - fis
13	260.5	14.5	294	2	A40444	protein kinase (EC
14	259.5	14.5	311	2	S36619	protein kinase (EC
15	259.5	14.5	382	2	JC5531	c-Jun amino-termin
16	259	14.5	391	2	I49141	casein kinase (EC
17	259	14.5	350	2	S20404	casein kinase II (
18	258.5	14.4	294	2	S22440	protein kinase (EC
19	258	14.4	391	2	JN0555	casein kinase II (
20	257.5	14.4	294	2	B40444	protein kinase (EC
21	257.5	14.4	278	2	S36387	probable protein k
22	256.5	14.3	826	3	JC5153	mitogen-activated
23	256	14.3	305	2	S23382	protein kinase (EC
24	255.5	14.3	297	1	TV2P2	protein kinase (EC
25	255.5	14.3	420	2	A43297	casein kinase II (
26	255.5	14.3	424	2	A56358	protein kinase (EC
27	255.5	14.3	382	2	S71100	protein kinase JNK
28	255.5	14.3	435	2	S64950	protein kinase HOG
29	255	14.2	288	2	S42566	protein kinase (EC
30	255	14.2	369	2	S56638	mitogen-activated
31	254.5	14.2	302	2	I50474	protein kinase (EC
32	254.5	14.2	424	2	A55480	c-Jun amino-termin
33	252	14.1	346	2	I48157	protein kinase (EC
34	251.5	14.0	423	2	S43968	p54-alpha-2 stress
35	251	14.0	350	2	B38611	casein kinase II (
36	251	14.0	415	2	A56042	mitogen-activated
37	251	14.0	293	2	T02922	protein kinase (EC
38	250.5	14.0	302	2	B44349	protein kinase (EC
39	248.5	13.9	332	2	S44355	casein kinase II (

40 248.5 13.9 384 2 S71098 protein kinase JNK
41 248.5 13.9 427 2 S71099 protein kinase JNK
42 248 13.8 350 2 B55838 casein kinase II (protein kinase II (protein kinase (EC stress-activated p protein kinase (EC
43 248 13.8 424 2 SI3934
44 248 13.8 427 2 JC5694
45 247.5 13.8 294 2 JQ2243

ALIGNMENTS

RESULT 1
S68675
Mitogen-activated protein kinase (EC 2.7.1.-) - fission yeast (Schizosaccharomyces po
N:Alternate names: MAP kinase; protein SPAC24B11.06c
C:Species: Schizosaccharomyces pombe
C:Date: 25-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 16-Dec-1998
C:Accession: S68675; S57930; S62551; S68433
R:Kato Jr., T.; Okazaki, K.; Murakami, H.; Stettler, S.; Fantes, P.A.; Okayama, H.
FEBS Lett. 378, 207-212, 1996
A:Title: Stress signal, mediated by a Hog1-like MAP kinase, controls sexual developme
A:Reference number: S68675
A:Accession: S68675
A:Molecule type: DNA
A:Residues: 1-349 <KAT>
R:Millar, J.B.A.; Buck, V.; Wilkinson, M.G.
submitted to the EMBL Data Library, June 1995
A:Description: Pyp1 and Pyp2 PIPases dephosphorylate an osmosensing MAP kinase contro
A:Reference number: S57930
A:Accession: S57930
A:Molecule type: DNA
A:Residues: 1-349 <MIL>
A:Cross-references: EMBL:X89262; NID:9897809; PID:g897810
R:Odell, C.; Churcher, C.M.
submitted to the EMBL Data Library, November 1995
A:Reference number: S62546
A:Accession: S62551
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-349 <ODE>
A:Cross-references: EMBL:Z67757; NID:g1061288; PID:g1061294
R:Shiozaki, K.; Russell, P.
Nature 378, 739-743, 1995
A:Title: Cell-cycle control linked to extracellular environment by MAP kinase pathway
A:Reference number: S68433; MUID:96107317
A:Accession: S68433
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-349 <SHI>
A:Cross-references: GB:U26739; NID:g1022684; PID:g1022685
C:Genetics:
A:Gene: STY1; SPC1; PHH1
A:Map position: 1L
C:Superfamily: kinase-related transforming protein; protein kinase homology
C:Keywords: Arp; phosphoprotein; phosphotransferase; protein kinase
F:18-268/Domain: protein kinase homology <KIN>
F:26-34/Region: protein kinase ATP-binding motif
F:171/Binding site: phosphate (Thr) (covalent) #status predicted
F:173/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 15.4%; Score 276.5; DB 2; Length 349;
Best Local Similarity 26.9%; Pred. No. 1.4e-11;
Matches 93; Conservative 60; Mismatches 114; Indels 79; Gaps 15;
QY 2 KLSDDYYDKELIYNSAIDYYTAIDKFNPLPVCLKIVDEDFSLP--PHSHRIFILKTL 59
Db 15 EITTRYSDLOPGMGAFGLVCSAKDQLTGMNVAVKIMKPFSTPLAKRTYRELKIKHL 74
QY 60 KPHNIYEYENDLKI--YDDVILVTKLYRYDLSQLTEIKYCKRTTREFYINGNLVSNQ 117
Db 75 R-HENIIS-LSDFISFPFDIYFVTELLGDLHRL--TSR----- 112

[illegible]

```

Query Match      14.6%; Score 262; DB 2; Length 391;
Best Local Similarity 26.5%; Pred. No. 1.5e-10;
Matches 90; Conservative 59; Mismatches 125; Indels 66; Gaps 13;

QY  5 DYVDKELYNLSAIDYITADIKFNLPVCLKIVDEDFSLPPHSIHREIFILKTLPHPN 64
   || : : || : : || : : || : : || : : || : : || : : || : : ||
Db  38 DYQLVRLK-GRGKYSEVFRAINTNEKVVVKLK---PVKKKKIKREIKLENLHGPN 93

QY  65 IYEFNDKLKYDVLIVTLKLYRDLSQLIEITKYCKRTRFYIYINGNLVSNQYITLANEI 124
   || : : || : : || : : || : : || : : || : : || : : || : : ||
Db  94 IITLADIVK--DPV-----SRTPALVFEHVNTTDFKQ--LYQTL 128

QY  125 EEKDIKLWLKSSGLEFTHSGIIHRDKPSNIFARDDITOTPIIGDFDICVDLKLPPK 184
   || : : || : : || : : || : : || : : || : : || : : || : : ||
Db  129 TDYDIRFYWEIKALDYCHSGMIGMRDVKPHNMVIDEHRKURLI-----DWGLAEF 181

QY  185 DEPPMAKYIDVSTGYIKAPELILGIITNYEYIEDIWSLGIILTGLY--SENFQSVLVKDDK 242
   || : : || : : || : : || : : || : : || : : || : : || : : ||
Db  182 YHPGOEYNNRVASRYKPKPELLVDYQMYDSLDMWSLGCMLASMIKPEPF----- 233

QY  243 ELTNDSHVSDLY-LLNQIIFENGTPLNIDFEELFCODEYNNENLHFKKN--LQKYPRKD 299
   || : : || : : || : : || : : || : : || : : || : : || : : ||
Db  233 -----FHGNDVDQLVRIAKVLGTEDLYD-----YIDKYNIE--LDPREFNDILGRHSKR 280

QY  300 WDIIILPRCNDLMLK-----BIFTKMIYDRSKRITSKEILQ 335
   || : : || : : || : : || : : || : : || : : || : : || : : ||
Db  281 WERFVHSENOHLVSPALDFLDKLLRYDHQSRLTAREAME 320

RESULT 10
A41227
protein kinase (EC 2.7.1.37) cdk2 - human
N:Alternate names: Egi homolog; protein kinase p34
C:Species: Homo sapiens (man)
C:Date: 17-Jul-1992 #sequence revision 17-Jul-1992 #text_change 24-Sep-1998
C:Accession: A41227; S17873; S16520
R:Ninomiya-Tsuji, J.; Nomoto, S.; Yasuda, H.; Reed, S.I.; Matsumoto, K.
Proc. Natl. Acad. Sci. U.S.A. 88, 9006-9010, 1991
A:Title: Cloning of a human cDNA encoding a CDC2-related kinase by complementation of a
A:Reference number: A41227; MUID:92020980
A:Accession: A41227
A:Molecule type: mRNA
A:Residues: 1-298 <NIN>
A:Cross-references: GB:M68520; NID:g180177; PID:g180178
R:Tsal, L.H.; Harlow, E.; Meyerson, M.
Nature 353, 174-177, 1991
A:Title: Isolation of the human cdk2 gene that encodes the cyclin A- and adenovirus Ela
A:Reference number: S17873; MUID:91367262
A:Accession: S17873
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-298 <TSA>
R:Elledge, S.J.; Spottswood, M.R.
EMBO J. 10, 2653-2659, 1991
A:Title: A new human p34 protein kinase, CDK2, identified by complementation of a cdc28
A:Reference number: S16520; MUID:91330891
A:Accession: S16520
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-176, 'S', 178-298 <ELL>
A:Cross-references: EMBL:X61622; NID:g29848; PID:g29849
C:Genetics:
A:Gene: GDB:CDK2
A:Cross-references: GDB:128984; OMIM:116953
A:Map position: 12q13-12q13
C:Superfamily: kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; cell cycle control; mitosis; phosphoprotein; phosphotransferase; serine
F:2-255/Domain: protein kinase homology <KIN>
F:10-18/Region: protein kinase ATP-binding motif
F:14, 160/Binding site: phosphate (Thr) (covalent) #status predicted
F:15/Binding site: phosphate (Tyr) (covalent) #status predicted

```


QY	60	KPHNIIIEYFN----	DLKIYDDVILTKLYRYDLSQIETIKTKYKTRTRFIYINGNLVS	115
Db	78	H-HPNIVNLVDVICADGKLY----	LIFEYVDHDLKALE-----	KRGGAF 118
QY	116	NOYTLANEIEEKDKLWLKSMSSGLEFHSQGIHRDIKPSNIFFARDDITQPIIGDFDI	175	
Db	118	-----TGTTEKKIIVQLLEGISFCHRRHIVHRDLKPAFILVTTDNSVK--	IADFGL 166	
QY	176	CYDLKLPPKDEPPMAKII-DVSTGIYKAPELILGITNYEYIOWISLGIILTLGLYSENFQ	234	
Db	167	ARAFOI-----	PMHTYTHEVTVLWYRAPELLGEXHYTAVDWMWSIGCIFAEIAR--	G 217
QY	235	SVLVKDDKELTNDSHVSDLYLLNOIFENFGTP-----	NLTFDEDELFCDEYNNE 283	
Db	218	KVLFPGDSEIQ-----	LFEIFVLGTPMDAEGSWLGVSLLPYRDV-----	260
QY	284	NLHFKKFNLOKYPKRDWIIILPCNDLMLKEIFTKIMIRYDRSKRITSKEILQ	335	
Db	260	-----FPKWSGKPLTQVLTPLDGDV-DLLSQMLRYNPAERISAKAALQ	302	
RESULT 15				
JC5531				
c-Jun amino-terminal kinase (EC 2.7.1.-) 2 alpha1 - chicken				
N:Alternate names: stress-activated protein kinase				
C:Species: Gallus gallus (chicken)				
C>Date: 02-Sep-1997 #sequence_revision 05-Sep-1997 #text_change 10-Jul-1998				
C:Accession: JC5531				
R:Ishikawa, T.; Nakada-Moriya, Y.; Ando, C.; Tanda, N.; Nishida, S.; Minatogawa, Y.; Noh				
Biochem. Biophys. Res. Commun. 234, 489-492, 1997				
A:Title: Expression of the JNK2-alpha1 gene in the developing chick brain.				
A:Reference number: JC5531; MUID:97320448				
A:Accession: JC5531				
A:Status: nucleic acid sequence not shown				
A:Molecule type: DNA				
A:Residues: 1-382 <ISH>				
A:Cross-references: DDBJ:AB000807; NID:g1816447; PID:d1019937; PID:g1816448				
C:Comment: This enzyme is involved in development of the central nervous system as a mediator. It plays a role in the differentiation of central nervous system.				
C:Genetics:				
A:Gene: JNK2-alpha1				
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog				
C:Keywords: phosphotransferase				
F:24-278/Domain: protein kinase homology <KIN>				
F:230/Active site: His #status predicted				
Query Match 14.5%; Score 259.5; DB 2; Length 382;				
Best Local Similarity 27.9%; Pred. NO. 2.1e-10;				
Matches 95; Conservative 49; Mismatches 109; Indels 87; Gaps 17;				
QY	21	IYTAIDKFNFLPVCILKIYDEDFSLPHS--	IHREIFILTKLPHPNIIIEYFN-----	DL 72
Db	40	VCAAFDTVLGINVAVKLSRPFQNTAKRAYRELVLKCVN-HKNIISLNVFTTQKSL	98	
QY	73	KIYDDVILTKLYRYDLSQIETIKTKYKTRTRFIYINGNLVSQYTLANEIEEKDKILW	132	
Db	99	EEFQVYLVLMELMDANLCQVIM-----	ELDHKMSYL 131	
QY	133	LKSMSSGLEFHSQGIHRDIKPSNIFFARDDITQPIIGDFDI----	CYDLKLPPKDEPP 188	
Db	132	LYQMLCGIKHLHSAGIIHRDLKPSNI-VKSDCTLKIL-DFGLARTACTNFMTPY----	186	
QY	189	MAKYIDVSTGIYKAPELILGITNYEYIOWISLGIILTLGLYSENFOSVLVWDDKELTND	248	
Db	186	-----VVTRYRAPEVITLGM-GYKENVDIWSVGCIM-----	GELVKGCVIFQGTD 229	
QY	249	HVSDLYLLNOIFENFGTNPENLDFEDELFCDEYNENLHFKKFNLOKYPK-----	KDWD 301	
Db	230	HIDQ--WNKVIQLGTPS-AEFMKKLQPTVRNVE-----	NRPKYPGKFEELPDM- 279	
QY	302	IILPCND-DLMK-----	EFTKMYDRSKRITSKEILQ 335	

Db 279 -IFPESDRDKLKTQARDLLSKMLVVDPKRISVDEALR 317

Search completed: November 6, 1999, 08:29:57

Job time: 452 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: November 6, 1999, 02:20:53 ; Search time 438.25 Seconds
(without alignments)
7394.718 Million cell updates/sec

Title: US-09-072-994-13
Perfect score: 1019
Sequence: 1 ATGAAGTGTGCAGATTATTA.....TTACAATTAATGTTGGATTG 1019

Scoring table: IDENTITY_NUC

Searched: 679419 seqs, 1590154680 residues

Database :

GenEmbl.*

- 1: gb_ba1.*
- 2: gb_ba2.*
- 3: gb_ov.*
- 4: gb_ov.*
- 5: gb_pat.*
- 6: gb_ph.*
- 7: gb_pl1.*
- 8: gb_pl2.*
- 9: gb_pr1.*
- 10: gb_pr2.*
- 11: gb_pr3.*
- 12: gb_ro.*
- 13: gb_st.*
- 14: gb_sts.*
- 15: gb_sy.*
- 16: gb_un.*
- 17: gb_v1.*
- 18: em_fun.*
- 19: em_htg.*
- 20: em_hum1.*
- 21: em_hum2.*
- 22: em_in.*
- 23: em_om.*
- 24: em_or.*
- 25: em_ov.*
- 26: em_pat.*
- 27: em_ph.*
- 28: em_pl.*
- 29: em_ro.*
- 30: em_sts.*
- 31: em_sy.*
- 32: em_un.*
- 33: em_v1.*
- 34: gb_htg1.*
- 35: gb_htg2.*
- 36: gb_in1.*
- 37: gb_in2.*
- 38: em_ba1.*
- 39: em_ba2.*
- 40: em_hum3.*
- 41: em_hum4.*
- 42: gb_pr4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
c 1	93.2	9.1	168935	34	HS1108D11	AL034419 Homo sapi

2	89.6	8.8	187544	35	AC004157	AC004157 Plasmodiu
3	89.6	8.8	104992	35	AC005504	AC005504 Plasmodiu
4	88.8	8.7	3542	7	YSCMTG16	L36900 Saccharomyc
5	87.8	8.6	14867	37	AE001398	AE001398 Plasmodiu
6	86	8.4	14001	36	PFCOMP18B	X95276 P.falciparu
7	85.8	8.4	162996	42	AC006441	AC006441 Homo sapi
8	85.6	8.4	168935	34	HS1108D11	AL034419 Homo sapi
9	85.6	8.4	187544	35	AC004157	AC004157 Plasmodiu
10	85.6	8.4	104992	35	AC005504	AC005504 Plasmodiu
11	85.6	8.4	102353	35	AC007708	AC007708 Homo sapi
12	84.6	8.3	1867	7	MTSCAJ23	AJ223323 Saccharom
13	84.4	8.3	85779	7	SCF011856	AJ011856 Saccharom
14	83.6	8.2	86829	36	PFMAL3P5	U98222 Saccharomyc
15	83.4	8.2	2426	7	SDU49822	U98222 Saccharomyc
16	83.4	8.2	173839	35	AC007483	AC007483 Homo sapi
17	83	8.1	110000	34	PFMAL4P1.2	Continuation (3 of
18	82.6	8.1	15421	36	PFCOMP18A	X95275 P.falciparu
19	82.4	8.1	5371	7	YSCMTTGS	J01462 Yeast (S.ce
20	81.6	8.0	120610	34	AC004948	AC004948 Homo sapi
21	81.6	8.0	15421	36	PFCOMP18A	X95275 P.falciparu
22	81.4	8.0	14867	37	AE001398	AE001398 Plasmodiu
23	81.2	8.0	27694	7	HASMT	D31785 Hansenula w
24	81	7.9	176552	11	AC004617	AC004617 Homo sapi
25	81	7.9	173839	35	AC007483	AC007483 Homo sapi
26	80.8	7.9	86829	36	PFMAL3P5	AL034556 Plasmodiu
27	80	7.9	4601	36	DMU11584	U11584 Drosophila
28	80	7.9	19517	37	DMG37541	U37541 Drosophila
29	79.6	7.8	12940	37	AE001386	AE001386 Plasmodiu
30	79	7.8	16019	36	MIDYTRN	U11584 Drosophila
31	79	7.8	2290	36	MIDYTRN	X03240 Drosophila
32	79	7.8	13594	37	AE001367	X05915 D.yakuba m
33	79	7.8	19517	37	DMU37541	AE001367 Plasmodiu
34	79	7.8	162956	42	AC006441	U37541 Drosophila
35	78.8	7.7	1200	7	MISCVA	AC006441 Homo sapi
36	78.8	7.7	2305	36	DTMTRNA	V00705 Yeast mitoc
37	78.8	7.7	85779	7	SCF011856	X54011 D.teissleri
38	78.2	7.7	284972	35	AC005505	AJ011856 Saccharom
39	78.2	7.7	220475	35	AC006278	AC005505 Plasmodiu
40	78	7.7	231699	35	AC006281	AC006278 Plasmodiu
41	78	7.7	134743	11	HUAC002300	AC006281 Plasmodiu
42	77.8	7.6	27694	7	HASMT	AC003300 Homo sapi
43	77.2	7.6	176552	11	AC004617	D31785 Hansenula w
44	77.2	7.6	176552	11	AC004617	AC004617 Homo sapi
45	76.8	7.5	194410	35	AC005140	AC005140 Plasmodiu

ALIGNMENTS

RESULT 1	HS1108D11/c	HS1108D11	168935 bp	DNA	HTG	02-JUN-1999
LOCUS	Homo sapiens chromosome 20 clone 1108D11, WORKING DRAFT SEQUENCE,					
DEFINITION	Homo sapiens chromosome 20 clone 1108D11, WORKING DRAFT SEQUENCE,					
	in unordered pieces.					
ACCESSION	AL034419					
NID	94995265					
VERSION	AL034419.10	GI:4995265				
KEYWORDS	HTG; HTGS_PHASE1.					
SOURCE	human.					
ORGANISM	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;					
	Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1 (bases 1 to 168935)					
AUTHORS	Clark, G.					
TITLE	Direct Submission					
JOURNAL	Submitted (02-JUN-1999) Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk					
COMMENT	On Jun 3, 1999 this sequence version replaced gi:4756985. IMPORTANT: This sequence is unfinished and does not necessarily represent the correct sequence. Work on the sequence is in progress and the release of this data is based on the understanding that the sequence may change as work continues. The sequence may be					

ORIGIN

Query Match 8.8%; Score 89.6; DB 35; Length 187544;
Best Local Similarity 44.5%; Pred. No. 0.00034;
Matches 444; Conservative 0; Mismatches 549; Indels 5; Gaps 2;

QY 17 ATTATATAGACAAGAAATTAATTACAAATAGTGCATTTCTGATATATATACGGCTATTG 76
DB 169324 ATAATAAATAATTATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATT 169383
QY 77 ATAAGTTTAATAAATCACTACCACTATGCTCTTAATAATAGTTGATGAGATTTTCAGTCTCCAC 136
DB 169384 AAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 169443
QY 137 CACATTCATCAATCGAGAAATTTTATACATTAATAAATTTTGAACACCAATCCCAACATAA 196
DB 169444 AAATTTTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 169503
QY 197 TTGAATATTTAATGATCTTAATAATTTATGATGATGTATATTAGTCACCAATTTGATC 256
DB 169503 TAAATTAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 169563
QY 257 GTTATGATTGAGTCAATTTGAAATTTACAAATTTTGAACGACCAACACGATTTA 316
DB 169564 ATAATAATATATATTAAACAATTAATTTATTAATAAATAAATAAATAAATAAATAAATAA 169623
QY 317 TTTATGGTATTAAAGTAACTTGTAGTAATCAATATACATTTGCTAAATGAATTTGAAG 376
DB 169624 TAAATTAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 169683
QY 377 AAAAAATCAATCAATTTGTTTAAATCAATGAGTTCAGGACCTTGAATTTATTCATTCAC 436
DB 169684 TTATATATTAGTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 169743
QY 437 AAGGATATATCATCGTATATAAACCAGCAATATATTTCTTGGCCGGATGATATAA 496
DB 169744 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 169803
QY 497 CACAACCGATTTATGGAGATTTTGATATTTGCTATGATTTTAAACCTCCACCTAAAGATG 556
DB 169804 ATATATATACAAATTAATAATATATATATATATTAATAATTAATAATATATATATAA 169863
QY 557 AACCCCTATGCGGAATATATGATGTATCTACAGGTATTTTAAAGACCCAGATTTGA 616
DB 169864 TTAATAAT 169923
QY 617 TTTCTGGTATACATTAATTAATGAATATGAAT -TGATTTTGGTCAATGGGTATATATTTG 675
DB 169924 TAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 169983
QY 676 ACTGGTTTATATCAAGAAATTTTCAAGTGTGTTTGTAGTCAAGATGATGAAGATTTGACT 735
DB 169984 AATAAATTAAT 170043
QY 736 ATGATTTCTAGTTAGTATTTATATTTATTAATCAATATTTGAAATTTTGGGTACA 795
DB 170044 AATAAT 170103
QY 796 CCCAATTTAACTTGATTTGAAGATGAATTTTGTGATGAATATATATATATATATATATATAT 855
DB 170104 AATAATTAATAAT 170163
QY 856 CATTTTAAAAAATTCATTTTCAAAATATCCAGAAAGATTTGGGATATATATATATATATATAT 911
DB 170164 AATAATTAACCAATTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 170223
QY 912 ACTCGATGCAATGATGATTTAATGAAAGAAATTTTACCAGATGATTTAGATGATGATCG 971
DB 170224 AATAATTAACCAATTTAATTAATTTAGATATATATATATATATATATATATATATATATAT 170283
QY 972 AAGTAAAGAAATCTTCAAGAAATCTTACAATTA 1009

Db 170284 TAATAAATGTTTTATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 170321

RESULT 3

AC005504 104992 bp DNA HTG 01-APR-1999
Plasmodium falciparum chromosome 12, WORKING DRAFT SEQUENCE, 3
unordered pieces.
AC005504
NID 94558584
VERSION AC005504.3 GI:4558584
KEYWORDS HTG: HTGS PHASE1.
SOURCE malaria parasite P. falciparum.
ORGANISM Plasmodium falciparum
REFERENCE Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
AUTHORS Hymen, R.W., Fung, E.L., Qin, F., Tamaki, T., Kurdi, O.B., Conway, A.B.
and Davis, R.W.
TITLE Plasmodium falciparum 3D7 chromosome 12
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 104992)
TITLE Hymen, R.W., Qin, F., Fung, E.L., Conway, A.B. and Davis, R.W.
JOURNAL Direct Submission
COMMENT Submitted (21-AUG-1998) Stanford DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
On Apr 2, 1999 this sequence version replaced gi:4337172.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 58642: contig of 58642 bp in length
* 58643 58842: gap of unknown length
* 58843 91011: contig of 32169 bp in length
* 91012 91211: gap of unknown length
* 91212 104992: contig of 13781 bp in length.
FEATURES Location/Qualifiers
source 1..104992
/organism="Plasmodium falciparum"
/db_xref="taxon:5833"
/chromosome="12"

BASE COUNT 44286 a 9326 c 9564 g 4141 t 405 others
ORIGIN

Query Match 8.8%; Score 89.6; DB 35; Length 104992;
Best Local Similarity 44.5%; Pred. No. 0.00041;
Matches 444; Conservative 0; Mismatches 549; Indels 5; Gaps 2;

QY 17 ATTATATAGACAAGAAATTAATTACAAATAGTGCATTTCTGATATATATACGGCTATTG 76
DB 72789 ATATATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 72848
QY 77 ATAAGTTTAATAAATCACTACCACTATGCTTTAAATAGTTGATGAGATTTTCAGTCTCCAC 136
DB 72849 AAAATATAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 72908
QY 137 CACATTCATCAATCGAGAAATTTTATACATTTTAAACTTTGAAACCCACATCCCAACATAA 196
DB 72909 AAATTTTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 72968
QY 197 TTGAATATTTTAATGATCTTTAAATTTTATGATGATGTTTATTTAGTCCCAAAATTTGATC 256
DB 72969 TAAATTAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 73028
QY 257 GTTATGATTGAGTCAATTTGATTCAAATTTACAAATTTTGAACGACCAACGATTTA 316
DB 73029 AATAATAATATATTAATAAACAATTAATTTATTAATAAATAAATAAATAAATAAATAAATAA 73088

repeat_region 12282..12619
/rpt_family="MER18"
complement(12620..12649)
/rpt_family="MIR"
repeat_region 12656..12809
/rpt_family="L1MA2"
complement(12941..13659)
/rpt_family="LTR1B"
13662..13779
/rpt_family="L1MA3"
repeat_region 13780..14075
/rpt_family="AluJb"
14076..14240
/rpt_family="L1MA3"
complement(14333..14627)
/rpt_family="AluY"
repeat_region 14628..14651
/rpt_family="AT-rich"
complement(14691..14761)
/rpt_family="MIR"
repeat_region 15905..16202
/rpt_family="AluJo"
complement(16333..16724)
/rpt_family="MSTC"
repeat_region 16777..16993
/rpt_family="MIR"
complement(17020..17125)
/rpt_family="L1MC5"
repeat_region 17141..17428
/rpt_family="AluSg1"
complement(17432..17570)
/rpt_family="AluSg/x"
repeat_region 17582..17653
/rpt_family="L1MC5"
17954..18218
/rpt_family="AluSp"
repeat_region 18219..18289
/rpt_family="(GGGA)n"
18359..18405
/rpt_family="(TG)n"
repeat_region 18639..18676
/rpt_family="(TG)n"
18878..19179
/rpt_family="AluJb"
repeat_region 19202..19503
/rpt_family="AluSx"
complement(20063..20356)
/rpt_family="AluSx"
repeat_region 20835..20877
/rpt_family="L2"
20878..20952
/rpt_family="GA-rich"
complement(21223..21533)
/rpt_family="AluSx"
repeat_region 21991..22245
/rpt_family="L2"
complement(22402..22474)
/rpt_family="MIR"
repeat_region 22472..22790
/rpt_family="L2"
complement(22924..23524)
/rpt_family="L2"
23525..23829
/rpt_family="AluJb"
complement(23830..23861)
/rpt_family="L2"
23936..24244
/rpt_family="AluSx"
24247..24443
/rpt_family="AluSg/x"
25664..25971
/rpt_family="AluY"

Query Match 8.4%; Score 85.8; DB 42; Length 162996;
Best Local Similarity 45.7%; Pred. No. 0.0011;
Matches 418; Conservative 0; Mismatches 487; Indels 10; Gaps 3;
QY 33 ATTAATTTACAAATAGTCCCAATTTCTGATATATATACGGCTATTGATGAAGTTTAAATACATT 92
Db 78414 ATGTTTATATTTATATATATATTTTATATGTTTATATATATATATATATATATATATTT 78473
QY 93 ACCAGTATGCTTTAAATAGTTGATGAAGATTTCAGTCTTCCACCACATTCATCCATCG 152
Db 78474 A--TATATGTTTATATATCTATATATATTTATATGTTTATATATATTTTATATGTTTAT 78531
QY 153 AGAAATTTTATATCTTAAACATTTGAACCCACATCCAAACATAATGATATATTTTAAATGA 212
Db 78532 ATATATTTGATATATATATATATTTTATATATATATATATATATATATATATATATATAT 78591
QY 213 TCTTAAATTTATGATGATGTTATATATAGTCACCAATTTGATCGTATGATTTGAGTCA 272
Db 78592 ATTTATATATTCATATATATATATATATATTTTATATATTCATATATATATATATATTT 78651
QY 273 ATTGATTTGAATTTACAAATATTTGTAACGACACACAGATTATTTATGTTATTAATGG 332
Db 78652 ATATATTCATATATATTTTATATAGTTATATATATTCATATATATATATAGTTATATATTC 78711
QY 333 TAACTCTGTTAGTAAATCAATATACACTTCTGCTAATGAATTTGAAGAAAGATATCAAAATT 392
Db 78712 ATATATATTTATATAGTTATATATATATTCATATATATTTATATAGTTATATATATATAT 78771
QY 393 ATGGTTTAAATCAATGAGTTCAGGACTTGAATTTTATTCATTCACAAAGGAAATTCATCG 452
Db 78772 ATATTTATATAGTTATATATATATTCATATATATTTTATATAGTTATATATATATATATAT 78831
QY 453 TGATATTAACCCAGTATATATTTCTTTCGCGGGATGATATAACACACCCGATTTATGG 512
Db 78832 TTATATAGTTATATATATTCATATATATTTTATATAGTTATATATATATATATATATATAT 78891
QY 513 AGATTTTGATATTTGTTATGATTTTAAACCTGCCACCTAAAGATGAACCCCTATGGCGAA 572
Db 78892 ATATAGATATATATATTTTATATATATATATATATATTTTATATATATATATATATATAT 78951
QY 573 ATATATGATGATCTACAGGTATTTTATAAGCACCCAGAAATTTGATCTTGGTATAACTAA 632
Db 78952 ATATTTATATATTTTATATATATTTTATATATATTTTATATATATTTTATATATATATAT 79011
QY 633 TTATCAATATGAATTTGATATTTGGTCAATTTGGGTATATATTTTGGCTGTTTATATTCAGA 692
Db 79012 TTTTATATATATTTTATACATATTTTATATTTTATATATATATATTTT---GTATATATATA 79066
QY 693 AAATTTTCAAAGTGTTTTGTAGTCAAGATGATAAAGAAATTTGACTAATGATCTCATGTTAG 752
Db 79067 TATATTTTGTATATATTTTATATATATTTTATATATATTTTATATATATATATATATTTTATA--TAT 79123
QY 753 TGATTTATATTTTAAATCAAAATTTTGAATAATTTGGTACACCCCAATTTAACTGATTT 812
Db 79124 TTATATATTTTATGTTATTTATATATATTTTATATATATATATATATTTTATATGTTATTTAT 79183
QY 813 TGAAGATGAATTTTGTGATGAATATATTAATGAACAACTTGCATTTTAAAAAATTCAA 872
Db 79184 ATATAATATATATTTTATGTTATTTTATATATATATATATATATTTTATGTTATTTATATTTTA 79243
QY 873 TTTCAAAATATCTAGAAAAGATTTGGGATATATTTTACCTCGATGCAATGATGATTT 932
Db 79244 TATATATTTATATATTTTATATATTTTATATATTTTATATGTTATTTATATATATATATATAT 79303
QY 933 AATGAAGAAATTTT 947
Db 79304 TGTATATATATTTAT 79318

RESULT 8

HS1108D11

LOCUS

DEFINITION

HS1108D11 168935 bp DNA

Homo sapiens chromosome 20 clone 1108D11, WORKING DRAFT SEQUENCE,

HTG

02-JUN-1999

in unordered pieces.

ACCESSION	QUERY MATCH	SCORE	DB	LENGTH	INDELS	GAPS
NID	Best Local Similarity	45.4%	Pred. No. 0.0011			
VERSION	Matches 436; Conservative 0;	Mismatches 509;	Indels 15;	Gaps 3;		
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
COMMENT						
FEATURES						
Source						
BASE COUNT						
ORIGIN						

[illegible]


```
Db 964 TTATAAAATATATATATATATATATATATATATATAGTTCTTCTTACATAAAATATAA 905
QY 980 GAATAACTTCTAAAGAAATCTTACAAATTAAT 1010
Db 904 TAAAAATTATAAATATATTAATTAAT 874

RESULT 13
SC0011856
LOCUS
DEFINITION Saccharomyces cerevisiae complete mitochondrial genome.
ACCESSION AJ011856
NID 94160362
VERSION AJ011856.1 GI:4160362
KEYWORDS baker's yeast.
SOURCE Mitochondrion Saccharomyces cerevisiae
ORGANISM Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;
Saccharomycetaceae; Saccharomyces.
REFERENCE 1 (bases 1 to 85779)
AUTHORS Foury,F., Roganti,T., Lecrenier,N. and Purnelle,B.
TITLE The complete sequence of the mitochondrial genome of Saccharomyces
cerevisiae
JOURNAL FEBS Lett. 440 (3), 325-331 (1998)
MEDLINE 99087401
REFERENCE 2 (bases 1 to 85779)
AUTHORS MIPS.
TITLE Direct Submission
JOURNAL Submitted (16-DEC-1998) Data collected by MIPS on behalf of the
European yeast mitochondrial genome sequencing project. MIPS at the
Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a D-82152
Martinsried, Germany; E-mail: Mewes@mips.biochem.mpg.de
FEATURES
Source
1. 85779
/organism="Saccharomyces cerevisiae"
/mitochondrion
/strain="F11679, isogenic derivative of S288C"
/db_xref="taxon:4932"
1. 11
/note="9S ribosomal RNA"
731. 802
/gene="tp(TGG)Q - systematic name"
/note="tRNA-Pro - common name; anticodon gene: TGG"
731. 802
/gene="tp(TGG)Q - systematic name"
3952. 4338
/note="ORF Q0010"
/codon_start=1
/transl_table=3
/protein_id="CAA09821.1"
/db_xref="PID:e1370887"
/db_xref="PID:g4160363"
/db_xref="GI:4160363"
/translation="MYIMFLYMLLIILIFYSIVGVPIIFNNNYWDPDIFLFI
YYKFIIFNLVYIMINIVTPSGSPGRTYILLYNNLYSYNMFIDYVMKFTIC
VYMLFMWLLSPSPYVYSEVPVS"
complement(4012..4312)
/note="Orl1"
4254. 4415
/note="ORF Q0017"
/codon_start=1
/transl_table=3
/protein_id="CAA09822.1"
/db_xref="PID:e1370688"
/db_xref="PID:g4160364"
/db_xref="GI:4160364"
/translation="MCATYMFNTVITHTPTLTRGPGFVRNRDLIYYKSNLIN
NLNNMTYL"
6546. 8194
/note="15S ribosomal RNA"
9374. 9447
/gene="tw(TCA)Q - systematic name"
```

```
gene
CDS
/note="tRNA-Trp1 - common name; anticodon gene: TCA"
9374. 9447
/gene="tw(TCA)Q - systematic name"
11667. 11957
/note="ORF Q0032"
/codon_start=1
/transl_table=3
/protein_id="CAA09823.1"
/db_xref="PID:e1370689"
/db_xref="PID:g4160365"
/db_xref="GI:4160365"
/translation="MLMMYMLFTMMKTYPMLSYHMMSYHIMLYTITMMYMKYSTYMLW
LLKSYRIFIFITNNNNYNYVMTMLNNTLYYNNHIFTINKKFLYSL"
complement(12510..12780)
/note="Orl8"
13818. 126701
/gene="COX1"
13818. 23167
/gene="SCA15alpha"
13818. 21935
/gene="I-SCEII"
13818. 19996
/gene="I-SCEIII"
13818. 18830
/gene="SCA12"
13818. 16322
/gene="SCA11"
/db_xref="SGD:S0007261"
join(13818..13986,16435..16470,18954..18991,20508..20984,
21995..23167)
/gene="SCA15alpha"
/note="ORF Q0070"
/codon_start=1
/transl_table=3
/protein_id="CAA09829.1"
/db_xref="PID:e1370695"
/db_xref="PID:g4160371"
/db_xref="GI:4160371"
/translation="MVORWLYSTNAKDIAVLVFMIAIFSGMAGTAMSLIIRLEAAPG
SOYLHGNSQLFNVLVGHAVIMIFFLVPALIGFGNYLLPLMIGATATAPRNNTA
FWPLMGLVCLVTSLVESGAGTGVTPPLSSIOAHSGSPVDLAFALHLSISLL
GAINFIVTLNMRNTGMHKLPLFVMSIFITAFLLLSLPVLSAGITMLLLDRNFT
SFEVSGGDPILYELHLPWFPGHEVYIILIPFGIISHVSTYSKPVFGEISNVYA
MASIGLGLFVASHHMYIVGLDADTRAYTSATMIAIPTGKIFSWMLNBSFKDKK
NKKRLRNQKNNNNMKTYLNNNNMMNNYKGNLDYDIPRSNRNYIQPNINKE
LVVYGNLSESCVGPTTINIVKMGVGPNNILYIMTGILLTDGWDYTSKDLDKKTI
MEINCRFLKQSMHSEYLMYFMLLSHYCSYPRMKIAKVKGKSYNQLSEFTKSLPC
FTILRYMFGVRKIVPNNLYDLNYESLAHIMMCDGSEVGGGLYLNLQSFYTKELI
FIMNLKIKFNLCNLTLHKSNNKYTIYMRVESVKRLFPMIYKYILPSMYKFDIMLWQK
KYNMIN"
13818. 16322
/gene="SCA11"
/note="ORF Q0050"
/codon_start=1
/db_xref="SGD:S0007261"
/transl_table=3
/protein_id="CAA09825.1"
/db_xref="PID:e1370691"
/db_xref="PID:g4160367"
/db_xref="GI:4160367"
/translation="MVORWLYSTNAKDIAVLVFMIAIFSGMAGTAMSLIIRLEAAPG
SOYLHGNSQLFNAGPTSAIYISMTALVLMINRYLKHNTSVGANFTGTWACHTPTM
ISVGVKCYMVRLTNFTLOVFIRITISSYHLDVKKVQVLFYEVIRLWFIVLDSTGSVK
KMDTNTNKTGKTSSEGSTGSDGVDRVNTOMKRFNLNOVRYYSVNNNLMKMGD
NIELSDTSTDILEPEKLVMDNNEENNLLSINKNVMDMLAYNRNISKPGNN
TPGTTLETLDGMNMYLNKLSNELGTGKFKPKPMRWNLPKPGKSWPLSVGNPRDKI
VQVWRMILDTIEDKMSHSGFRANNSCOIAIWEVRNMEGGSWFVEVLKCFDT
ISHDLILKLRKRTISDKGFDLVYKLRAGYIDKGTYKPMGLDQPSGLSPILCN
VMTLVDNWLDEYINLYNKGKVKQHPHTYKKLSRMIAKAMFSTRLKHREKAGPTFI
YNDENFRKMYVRYADDILIGVLGSKNDCKMIKRDNLNPNLSLGTMMNEETLITCAT
ETPAFLGYNISITPKRMPTVTKTIRGTIRSNRTPINAPIRDIRINLATNGYC
KHNKNGRMGVTRVGRWVEEPTIINNYKALGRGLNYKLATYKRLRRIYVLY
YSCVLTASKYRLKTSKTIKFGYNLNIENDKLIANFPNTFDNIKKIENHGMWY
```



```
gene      27547..28290
/ gene="MAL3P5.7"
CDS       27547..28290
/ gene="MAL3P5.7"
/ note="predicted using hexon; MAL3P5.7 (PFC0605c),
Hypothetical protein, len: 248 aa"
/codon_start=1
/protein_id="CAB41709.1"
/db_xref="PID:el427934"
/db_xref="PID:94725991"
/db_xref="GI:4725991"
/translation="MGSGGLNLLPQKKWVYRDAQYKVYDEHKLIKKEKDEIKR
KDEPSTISLKKWTKNEDSDNNYNFYDENGKKTINYCNHNLIDSEKELT
AKQKHEELIKGHYIYDANKFNTOHNSIYDKNKAQIISDFNKLCDERWFLNK
NKKETKDNGANFFHIQDNISEHNKNTENISDLSNTNNYIITHDKRKKRQMHY
HIKKIKYQEKDEKCKRKKRQKKEKKPK"
complement(29992..33537)
/ gene="MAL3P5.8"
complement(29992..33537)
/ gene="MAL3P5.8"
/ note="predicted using hexon; MAL3P5.8 (PFC0610c),
Hypothetical protein, len: 1182 aa"
/codon_start=1
/protein_id="CAB38971.1"
/db_xref="PID:el407791"
/db_xref="PID:94493935"
/db_xref="GI:4493935"
/translation="WAHVKKKEKTEAQETPVAKQETHAKENNESNIATVENVIS
KNGEIAISNDQEIATSKNDQEIATSKNDQENVALNSESERONASKEE
DNELKQIEFDHISNEHNEHNSFSTLSFFKEVNESVEQHFSEGHTEHSNE
DSNNVETIENLVLSNRSTYSKOKNELTSVTCVCGEYDLNWDHIFAHKL
DSKTRTFNSEKGNANVDCCKEDNLCVDYDEKKNINITNWSLSSSSNGAY
DSLNELYRINDEROKNKIAQPKEDDNNLTNSNTDNDTGTNTYDNRNLER
ENLEQEMYPNSKIDTCGYNTSLNGNETKIKESNNVNEKNSSEINLPSTENK
EMETLSRNSNNHNDKKEKRYNEKEETORSRYCDDNDSPINRYGAFNN
RKEFTIRDNPNSLHNDRTSYRIYRGINTSSSYNILENFYKRLRLQONRMYI
```

Query Match 8.2%; Score 83.6; DB 36; Length 86829;

Best Local Similarity 46.8%; Pred. No. 0.0024;

Matches 466; Conservative 0; Mismatches 519; Indels 11; Gaps 6;

```
QY 12 AGATTATATATACACAGGAATTAATTTACATATGCGCATTCGTGATATATATACGGC 71
DB 41213 ATATTATAAATACAAATAAATAATTTTCATAAATAATTAATAATATATAAAT 41154

QY 72 TATTGATAAGTTTAATAACTTACCAGTATGCTCTTAAATAGTTGATGAGATTTTCAGTCT 131
DB 41153 TAAATTAATATATATTAATTAATTAAGTGTATATATATATATATATAAATAAT 41094

QY 132 TCCACCACATCAATCCATCGAATAATTTTATACATAAACTTTGAACACCATCCAAA 191
DB 41093 TAAATATATAAATAATAAATAAATAAATTTTCATAAATAATTAATAATACATAAT 41034

QY 192 CATAAATTGAATATTTTAATCATCTTAAATTTTATGATGATGATATATAGTCACCAAAAT 251
DB 41033 AATAATTTTAATAAATAATTAATTAATTAATTAATTAATTAATTAATAAATAA 40974

QY 252 GTATCGTATGATTTAGTCAATGATGAAATACAAATAATTTGAACGACCAACAGC 311
DB 40973 TTA--ATTATATATAATTAAGATTTTAATTAATTAATTAATTAATTAATAAATAA 40916

QY 312 ATTTATTTTGGTAAATGATATCTGTGTAGTATCAATATACATGCTCAATGAAT 371
DB 40915 A-TTATATAATTAATAAATTAATATATATTAATTAATTAATAAATAAATAAATA 40857

QY 372 TGAAGAAAAGATATCAAAATATGTTTAAATCAATGAGTTCAGGACTTGAATTTATCA 431
DB 40856 TAAATTAATTAATAATAAATAATGTTATATAATTAATTAATTAATTAATTAATAA 40797

QY 432 TTCACAGGGATATTTCAATCGTGNATATAAACCCAGTATATTTCTTTGCCGGGATCA 491
DB 40796 TTAATAAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 40737
```

```
QY 492 TATAACACAAACCGATATTGGAGATTTTGATATTTGTTATGATTTAAACGCCACCTAA 551
DB 40736 TAAATATAATAAATAATATATATAATAATAATAATAATAATAATAATAATAATAA 40677

QY 552 AGATGAACCCCTATATGGCGAAATATATTGATGATCTACAGGTATTAT--AAAGCACCA 609
DB 40676 AATGTAATAAATAATATATATAATAATAATAATAATAATAATAATAATAATAA 40617

QY 610 GAATTGATTTCTTTGGTAACTAATATGAAATGAAATGATTTGGTCAATGGGTATA 669
DB 40616 AAATAATAAATAAATAATATATAATAATAATAATAATAATAATAATAATAATAA 40557

QY 670 ATTTTGACTGGTTTATATTCAGAAATTTTCAAGTGTTTCTAGTCAAGATGATAAGAA 729
DB 40556 AATTAATAAATAAATACTTT--AAATAAATAATAATAATAATAATAATAATAATAA 40500

QY 730 TTGACTAATGATTCCTCTAGTGTAGTATATATTTTAAATCAATATTTGAAATTTTC 789
DB 40499 TAAATACITTAATAAGTATATAATAATAATAATAATAATAATAATAATAATAA 40440

QY 790 GGTACACCAATTTAACTGATTTTGAAGATGAATTTTGTGATGATATATAAATGAA 849
DB 40439 TAAATAATTTTAAATTTTAAATTAATTAATTAATTAATTAATTAATAATAATAA -AA 40382

QY 850 AACTTCGATTTTAAATAATTTCAATTTTACAAAAATATCCTAGAAAAGATTTGGATATTT 909
DB 40381 TAATTAATTAATTAATTAATTAATAATAATAATAATAATAATAATAATAATAA 40322

QY 910 TTACCTCGATGCAATGATGATTTTAAATGAAGAAATTTT-TACCAAGATGATTAATGA 968
DB 40321 TTATATATTTTAAATAAATAAATAAATAATAATAATAATAATAATAATAATAATAA 40262

QY 969 TCGAAGTAAAAAATAAATCACTTCTAAAGAAATCTTACA 1004
DB 40261 TAATAATTAATTAATTTTAAATTAATAATAATAATAA 40226
```

RESULT 15

SDU49822

LOCUS

DEFINITION

ACCESSION

NID

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

SDU49822 2426 bp DNA PLN
Saccharomyces douglasii mitochondrial trna-ser and trna-phe genes.
partial sequence, and Varlp (varl) gene, mitochondrial gene
encoding mitochondrial protein, complete cds.

U49822 91236920
U49822.1 GI:1236920
mitochondrial trnser; varl; mitochondrial trnaphe.
baker's yeast.

Mitochondrion Saccharomyces douglasii
Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;
Saccharomycetaceae; Saccharomycetes.

1 (bases 1 to 1655)
Cardazzo, B., Pellizzari, R. and Carignani, G.
Characterization of the mitochondrial genome of S. douglasii:
analysis of the varl gene

Unpublished
2 (bases 1655 to 2426)
Cardazzo, B., Rinaldi, F., Frontali, L., Carignani, G. and Palleschi, C.
Evolution of the mitochondrial genome in yeast: a study of the
mitochondrial divergence in the two closely related species
S. douglasii and S. cerevisiae

Unpublished
3 (bases 1 to 2426)
Carignani, G.
Direct Submission

Submitted (22-FEB-1996) Giovanna Carignani, Università di Padova,
Chimica Biologica, via Trieste 75, Padova, Italy, 35121
Location/Qualifiers

1. 2426
/organism="Saccharomyces douglasii"
/mitochondrion
/strain="KSI58/3-2/BIE2"
/db_xref="taxon:46617"

```
tRNA      <!. .79
          /note="aa: Ser"
          /product="tRNA-serine"
gene       537..1658
          /gene="var1"
          537..1658
          /gene="var1"
          /function="mitochondrial ribosomal protein"
          /codon_start=1
          /transl_table=3
          /product="Var1p"
          /protein_id="AAA96260.1"
          /db_xref="PID:gl236921"
          /db_xref="GI:1236921"
          /translation="MKRLRLNMLSMNKTNNNNNNNNKLLKNNLMDNNKRN
          NKMTMLKNNNNINNKLOHLNMMNNTQIYNKNKMEIMNMNDKLNKLYIKMTIL
          KLNMMINKILMSKTIHSLNKLKIFYNNNDINNMMNNNNNNYNNMMNKLMMIM
          NNMNNSICNLSYTKNKKVTESIKLVIYLSNDFISKYISLNDMMKNNGILNTYO
          RMLNIMPKLNDHNTSMYINNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNYIDN
          INNIYNNMTIDNIPDILMYKYLVGSKIKEKGLNNNGRTSTINLLNGTFNNKKYLW
          SNINNKKUNYIPSNHNLYNNINNKNGKYNKIKVKNFI"
misc_signal 2201..2212
          /note="mRNA processing signal"
misc_signal 2247..2258
          /note="mRNA processing signal"
promoter    2345..2353
          /note="aa: Phe"
          2373..2426
          /product="tRNA-phenylalanine"
          /note="aa: Phe"
BASE COUNT 1184 a 100 c 106 g 1036 t
ORIGIN
```

Query Match		8.28;	Score 83.4;	DB 7;	Length 2426;
Best Local Similarity		45.5%;	Pred. No. 0.0083;		
Matches 414;		Conservative 0;	Mismatches 486;	Indels 9;	Gaps 3;
QY	1	ATGAAGTTGTCAGATTATTATATAGACAAGAAATTAATTTACAAATAGTGCATTTCTGAT	60		
Db	537	ATGAATTAAGATTATTAAATATATATTTTCAATATAATAAATAACTTTAATAATAAT	596		
QY	61	ATATATACGGCTATGTATAGTTTAAATACCTACCACTGTCTTAAATAGTTGATGAA	120		
Db	597	AATAATAATATTAAATAAGAAATTTATTATAAGAATATATTATTAGATATAAAAT	656		
QY	121	GATTTCAGTCTCCACACATCTCAATCCACGAGAAATTTTATACCTTAAACCTTTGAAA	180		
Db	657	AAAAGAATAAATAATAAATAAACAATATTAAATAATATAATAATAATAATAATAA	716		
QY	181	CCACATCCAAACATAATTTAATGATCTTAAATTTTAAATTTATGATGATGTTATATTA	240		
Db	717	TTACACACATTTAATAATATAAATAATTTGAATCTACAAATTTTATTAATAATAAAAT	776		
QY	241	GTCCACAAATTTGATCGTTATGTTGAGTCAATGATGAAATTCACAAATATTGTGAAA	300		
Db	777	ATAGAGATTATAAATATTATAAAT---GATAAATTAATTAATAAATTTATTATAAATA	833		
QY	301	CGAACACACAGATTATTATTGGTATTAAATCGTATCTGTTAGTAAATCAATATACACTT	360		
Db	834	ATACATTTAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA	893		
QY	361	GCTAATGAAATTTGAAGAAAAAGATACAAATTTATGGTTTAAATCAATGAGTTCAGGACTT	420		
Db	894	CAACATAGTTTAAATAAATAATAATAATAATAATAATAATAATAATAATAATAATAA	949		
QY	421	GAATTTATTCATCCAGGGGATATTCATCGTGNATATAAACCCAGTAATATTTCTTT	480		
Db	950	TAATAATAATAATAATAA--TAATAATAATTTATTAATAATAATAATAATAATAATAA	1007		
QY	481	CCCCGGGATGATATAACACACCGATTATTGGAGATTTTGGATTTTGTATGATGATTTAAA	540		
Db	1008	AATATTATAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA	1067		

Search completed: November 6, 1999, 07:48:50
Job time: 19677 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 6, 1999, 02:12:43 ; Search time 288.18 Seconds
(without alignments)
5974.848 Million cell updates/sec

Title: US-09-072-994-13
Perfect score: 1019
Sequence: 1 ATCAAGTTCGAGATTATTA.....TTACAATTAACTTGGATTG 1019

Scoring table: IDENTITY_NUC

Searched: 2546578 seqs, 986266752 residues

Database : EST.*

- 1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*
44: gb_est25:*
45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*

54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	54	5.3	610	44	AU038896	AU038896 AU038896
C 2	53.8	5.3	519	40	C91042	C91042 C91042 Dict
C 3	53.6	5.3	488	45	AU000986	AU000986 AU000986
C 4	53.4	5.2	705	43	A1238819	A1238819 GH14939.5
C 5	52.6	5.2	1145	48	A1581441	A1581441 SMOV3MCAM
6	52.4	5.1	684	44	A1294319	A1294319 LP07659.5
7	52.4	5.1	590	44	AU037557	AU037557 AU037557
8	52.2	5.1	478	43	AU033909	AU033909 AU033909
9	52	5.1	424	40	C92817	C92817 C92817 Dict
10	51.8	5.1	755	35	AA550575	AA550575 1740m3 gm
11	51.8	5.1	712	49	AU056862	AU056862 AU056862
12	51.6	5.1	579	44	AU038989	AU038989 AU038989
13	51.4	5.0	717	45	AU000694	AU000694 AU000694
14	51.4	5.0	688	45	AU001063	AU001063 AU001063
15	51.4	5.0	593	45	AU002045	AU002045 AU002045
16	51.4	5.0	633	45	AU004482	AU004482 AU004482
17	51.4	5.0	700	45	AU006427	AU006427 AU006427
C 18	51.2	5.0	717	45	AU000694	AU000694 AU000694
C 19	51	5.0	540	35	AA550445	AA550445 1598m3 gm
C 20	50.8	5.0	450	49	AI629364	AI629364 fc05e12 y
21	50.6	5.0	539	38	AA440967	AA440967 LD13786.5
22	50.4	4.9	594	38	AA539867	AA539867 LD13149.5
C 23	50.4	4.9	688	45	AU001063	AU001063 AU001063
24	50.4	4.9	736	45	AU001091	AU001091 AU001091
C 25	50.4	4.9	593	45	AU002045	AU002045 AU002045
C 26	50.4	4.9	633	45	AU004482	AU004482 AU004482
C 27	50.4	4.9	700	45	AU006427	AU006427 AU006427
C 28	50.2	4.9	428	35	AA549948	AA549948 1025m3 gm
C 29	50.2	4.9	578	35	AA550379	AA550379 1525m3 gm
C 30	50.2	4.9	644	40	AA942264	AA942264 LD26372.5
31	50	4.9	541	50	AU062000	AU062000 AU062000
32	49.8	4.9	514	20	T18187	T18187 0794c3 czap
33	49.8	4.9	536	38	AA440021	AA440021 LD15090.5
34	49.8	4.9	654	45	AU000948	AU000948 AU000948
C 35	49.8	4.9	551	45	AU000956	AU000956 AU000956
C 36	49.8	4.9	590	45	AU000995	AU000995 AU000995
C 37	49.8	4.9	573	45	AU001034	AU001034 AU001034
38	49.6	4.9	696	45	AU000518	AU000518 AU000518
39	49.4	4.8	529	26	W52352	W52352 zc48h03.s1
40	49.4	4.8	531	40	C90097	C90097 C90097 Dict
41	49.4	4.8	505	40	C90102	C90102 C90102 Dict
42	49.4	4.8	636	40	C92986	C92986 C92986 Dict
43	49.4	4.8	724	42	AI108800	AI108800 GH06024.5
44	49.4	4.8	681	44	AI260725	AI260725 LP04686.5
45	49.4	4.8	697	44	AI297514	AI297514 LP11826.5

ALIGNMENTS

RESULT 1
AU038896/c
LOCUS AU038896 Dictyostellium discoideum SS (H.Urushihara) Dictyostellium
DEFINITION AU038896 Dictyostellium discoideum SS (H.Urushihara) Dictyostellium
ACCESSION AU038896
NID 93965649
VERSION AU038896.1 Gr:3985649

610 bp mRNA
29-MAR-1999

COMMENT On Oct 6, 1998 this sequence version replaced gi:3712292.

Contact: Mita K
Genome Research Group
National Institute of Radiological Sciences
Anagawa 4-9-1, Inage, Chiba 263-8555, Japan
Email: knitaeuex64.nirs.go.jp
PROJECT = 'CREST project by JST'.

FEATURES

Location/Qualifiers
source
1..488
/organism="Bombyx mori"
/strain="p50(Daizo)"
/db_xref="taxon:7091"
/clone="fbf0818f"
/clone_lib="Bombyx mori p50(Daizo)"
211 a 15 c 31 g 231 t

BASE COUNT

211 a 15 c 31 g 231 t

Query Match 5.3%; Score 53.6; DB 45; Length 488;
Best Local Similarity 50.6%; Pred. No. 4.1;
Matches 157; Conservative 0; Mismatches 149; Indels 4; Gaps 1;

QY 594 TATTATATAGACACCAAGTTCCTGGTATCACTAATATGAATGAAATGTATAT 653
|||||
Db 432 TTTTATATATAAATTAATAAATTAATTTATCTATTAGTAAATTTATATAATA 373
QY 654 TTGTCATCGGTATATATTTGACTGTTTATATTCAGAAATTTTCAAGTCTTTT 713
|||||
Db 372 AATAAATTTATTAACATATAATTAATAATTTCTTTTAAATAAATTTATATAA 313
QY 714 CAAGATGATAAGAAATGTAATCTCAATGATCTAGTATTTATTTATTAATCA 773
|||||
Db 312 ATTTTAAATTTAAGATATATCCCTAAATTTATTTAATAAATATATTTTATTAATTA 253
QY 774 AATATTTGAAATTTGCGTACCCATTTAAGTATTTGAGATGAATTTTGTGA 833
|||||
Db 252 ATTAATAAATAATCTAATTAATAAATAAATTTTCTTATATAAATACTA---GATA 197
QY 834 TGAATATAATTAAGAAATCTGCATTTTAAAAAATCAATTTCAAAAAATATCTAGAAA 893
|||||
Db 196 TATTAAAAACGATAACATTTTCATTTCAATTAATTTTAAATATTTTATTAACAAT 137
QY 894 AGATTGGAT 903
|||||
Db 136 AGCTTATTAT 127

RESULT 4

AI238819 705 bp mRNA EST 18-NOV-1998
LOCUS GH14939.5prime GH Drosophila melanogaster head pot2 Drosophila
DEFINITION melanogaster cDNA clone GH14939 5prime similar to U37541;
Drosophila melanogaster complete mitochondrial genome, mRNA
sequence.

ACCESSION AI238819
NID 93833677
VERSION AI238819.1 GI:3833677
KEYWORDS EST.
SOURCE fruit fly.

ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 705)

REFERENCE Harvey, D., Hong, L., Evans-Holm, M., Pendleton, J., Su, C.,
AUTHORS Brokstein, P., Lewis, S. and Rubin, G.M.
BDGP/HMI Drosophila EST Project
TITLE Unpublished (1997)

COMMENT On Aug 21, 1998 this sequence version replaced.

Contact: Harvey, D.
G. M. Rubin-Molecular and Cell Biology

University of California Berkeley
539 LSA, Berkeley, CA 94720-3200, USA
Fax: 510 643 9947

Email: <http://www.fruitfly.org/EST>, est@fruitfly.berkeley.edu
Plate: 149 row: D column: 3
High quality sequence stop: 582.

FEATURES

Location/Qualifiers
source
1..705
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/map="959G05; 1; 17q21"
/clone="GH14939"
/clone_lib="GH Drosophila melanogaster head pot2"
/sex="male and female"
/dev_stage="adult"
/lab_host="DHS - alpha"
/note="Organ: head; Vector: pOT2; Site:1; EcoRI; Site:2;
XhoI; Sized fractionated cDNAs were directly ligated into
pOT2. Plasmid cDNA library."
315 a 26 c 58 g 306 t

BASE COUNT

315 a 26 c 58 g 306 t

Query Match 5.2%; Score 53.4; DB 43; Length 705;
Best Local Similarity 46.3%; Pred. No. 3.9;
Matches 285; Conservative 0; Mismatches 321; Indels 9; Gaps 3;

QY 316 ATTTATGTTTAAATGTTAACTCTTTAGTAAATCAATCAATCACTTGCTAATGAAATGAA 375
|||||
Db 34 ATTTAAAGTTTTAGTATTGTTTAAAGAAAAATAATTTTAAATAATAGTTTATTAGTATTG 93
QY 376 GAAAGATATCAATATATGTTTAAATCAATGAGTTCAGGACTTGAATTTATTCATCA 435
|||||
Db 94 TAAAGAAAAATGAAATTTGAAAAATTTTATTTAAAGAAAAATTTTATTTATTTATTTGT 153
QY 436 CAAGGGATAATTCATCGTGATATAAAACCCAGTAATATTTCTTTGCCGGATGATATA 495
|||||
Db 154 ACCTTGTTCAGTGTATTATAATAAATAAATAATTTATTTCTCGATTTTAAAG 213
QY 496 ACACAACGGATTTTGGAGATTTTG-----ATATTTGTTATGATTTTAAACTGCCCTA 550
|||||
Db 214 AGTTAATAATATATTAAGTTAATGTAATAATGTAATAATTTTATATAATATTATAGAAATGA 273
QY 551 AAGATGAACCCCTATGCGAAATATATTCATGCTATCTACAGGTATTTATAAAGCCACG 610
|||||
Db 274 AATGTTATTCGTTTTTAAAGTATCTAGTTTTTTAAGAAATAAATTTAATTTAGAA--AT 331
QY 611 AATTGATCTCTGTPATAACTAATTAATGAATATGAATATGAAATTTGATATTTGGTCATTTGGGTATAA 670
|||||
Db 332 TATAAATTTACTTAATTTATTTTAAATTAATTAATTTAATTTAATTTTATGTTTATGGA 391
QY 671 TTTTGAATGTTTATATTCAGAAAAATTTTCAAGTGTTTTAGTCAAGATGATAAAGAT 730
|||||
Db 392 TAAGCTATAAATAAATTTTAAATAATTTAATAATAGATTATAAATAATATATGCTTAGAAT 451
QY 731 TGACTAATGATCTCTGTTAGTGAUTTTATATTTTATTAATCAAAATTTTGAATAATTCG 790
|||||
Db 452 TAGCAATTTAATAAATAATGTTATATTTTATTTAATAATTAATTTTATTTAATTT 511
QY 791 GTACACCAATTTAACTG--ATTTGAAGATGAATTTTGTGATGAATATAATTAATGA 848
|||||
Db 512 TAATTTATTTAATAATGTTAGTTTTAATTTAATAATTTAATAATTAAGTAATATGATAAATAG 571
QY 849 AAAGTTGCAATTTAAAAAATTCATTTTACAAAAATATCTAGAAAAAGATTTGGATATTAT 908
|||||
Db 572 TATAATAAATGTTAAATAATTTTATATGAAAAAGTTTAAATAAAGATTTGCAAAAAATAA 631
QY 909 TTTACCTCGATGCAA 923
|||||
Db 632 TATTCGCTGTTTAA 646

RESULT 5

```

AI581441/c 1145 bp mRNA EST 06-APR-1999
LOCUS SMOV3MCAW25C06SK Onchocerca volvulus molting L3 larva cDNA
DEFINITION (SL96MLW-Ovml3) Onchocerca volvulus cDNA clone SMOV3MCAW25C06 5',
mRNA sequence.
ACCESSION AI581441
NID 94567335
VERSION AI581441.1 GI:4567336
KEYWORDS Onchocerca volvulus.
SOURCE Onchocerca volvulus.
ORGANISM Eukaryota; Metazoa; Nematoda; Secernentea; Spirurida;
Onchocerca volvulus.
REFERENCE 1. (bases 1 to 1145)
AUTHORS Williams, S.A., Lizotte-Waniewski, M., Lane, S. and Lustigman, S.
TITLE Genes expressed in molting L3 larvae of Onchocerca volvulus
JOURNAL Unpublished (1997)
COMMENT On Mar 20, 1998 this sequence version replaced gi:2980257.
Contact: Steven A. Williams
Molecular Parasitology
Smith College Department of Biological Sciences
Department of Biological Sciences, Clark Science Center, Smith
College, Northampton, MA, 01063, USA
Tel: 4135853826
Fax: 4135853786
Email: genomes@smith.edu
Seq primer: pluescript SK.
FEATURES
    source
        1..1145
            /organism="Onchocerca volvulus"
            /strain="Kumba, Cameroon"
            /db_xref="taxon:6282"
            /map="16"
            /clone="SMOV3MCAW25C06"
            /clone_lib="Onchocerca volvulus molting L3 larva cDNA
            (SL96MLW-Ovml3)"
            /dev_stage="molting L3"
            /lab_host="XLI-Blue MRF"
            /note="Vector: Lambda Uni-ZAP XR; Site_1: Eco RI; Site_2:
            Xho I; Filarial nematode parasite of humans. Third-stage
            larvae, L3, were isolated from infected black flies in
            Cameroon (forest strain). The L3 were cultured in 20% FCS
            in IMDM+ NCTC 135 and collected after day 1, 2, or 3 in
            culture. L3 of O. volvulus molt to fourth-stage larvae by
            day 5 in culture. mRNA was isolated from approximately
            6000 molting larvae (mL3), 2000 larvae from day 1, 2 or 3
            in culture, and converted to double-stranded cDNA using
            reverse transcriptase and oligo(dT) followed by RNase H
            and DNA pol I. The library was constructed in the lambda
            Uni-Zap XR vector and has 1 x 10E6 independent
            recombinants and the average insert size is ~1200 bp. The
            library was constructed by Sara Lustigman and Michelle
            Lizotte-Waniewski in the Laboratory of Dr. S. A. Williams.
            The library is available from Dr. Sara Lustigman (email:
            slustig@nyc.org)."
BASE COUNT 320 a 106 c 77 g 468 t 174 others
ORIGIN
Query Match 5.2%; Score 52.6; DB 48; Length 1145;
Best Local Similarity 39.0%; Pred. No. 4.6;
Matches 270; Conservative 0; Mismatches 412; Indels 11; Gaps 3;
QY 324 TATTATGGTAACTTGTAGTAATCAATATACACTTGGTATGAATGAGAAAGA 383
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1138 TATTAGGNTNATTNGTANATNCANCAATTAAGANGTNCANNGNANNANTTAAGA 1079
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 384 TATCAATATATGTTAAAAATCATGAGTTCAGGACTTGGAATTTATTCACAGGGAT 443
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1078 NATAGATAAAAAATTAATTATNGATNANTAAATGTAATACATAAAGAAATCGANAGNAA 1019
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 444 AATTCATCGTGATATAAACCCAGTAATTTTCITTCGCCGGGATGATATAACACACC 503
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

Db 1018 ANTTATNATTTTAAAGTAATATATATCAATTCATNGCAAAANNAATGANTAGTACGT 959
QY 504 GATTATTGGAGATTTTGTATTTTATTTGTTATGATTAA-AACTGCCACCTAAAGATGAACCCC 562
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 958 TAANCNGTATATATAAATATATATATATATATATATANTGTGTCATATATANGTGTAAATN 899
QY 563 CTATGCGGAAATATATGATGATCTACAGGTATTTTAAAGCACCAGAAATGATTTCTTG 622
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 898 ATANTATGAATATATATANTTTATGTTNATGANAACAANNANTNACATGATATANATAT-- 840
QY 623 GTATAACTAATATGAATGAATGATATTTGTCATTTGGGTATATATTTGACTGGTT 682
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 840 -NATCGTNNTCANATANGATAANAGATNGTNGAGAGATANACATATNTATNTAAGAA 782
QY 683 TATATTCAGAAATTTTCAAAGTGTGTTTAGTCAAGATGATAAAGAAATGACTTAATGATT 742
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 781 TATNNAATNTATNNAAAATNATNTANAANTATATATATATATATATATATGATNATAANTTAN 722
QY 743 CTCATGTTAGTGATTTATATTTTAAATCAAAATATTTGAAAATTTTCGGTACACCCAAAT 802
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 721 ATATTATGTAANATTAATTAATTTTAAANNATTANTATAATNTNNAANNAAN 562
QY 803 TAACGATTTTGAAGATGAATTTTTCGTGATGAATATAA-----TAATGAAACTTG 855
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 661 TNATTANNNTANNATTNAAANATTATNNGTANANATACAGTAAAGATANTAAAGATT 602
QY 856 CATTTTAAAAAATTCATTTACAAAAATATCCTAGAAAAGATTCGGATATATTTTACCT 915
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 601 ANTTTANATAATTAATTCANAAANTNCAATNAAATTTTANTATNTANATAAANTA 542
QY 916 CGATCGAATGATGATTTAATGAAGAAATTTTACCAGATGATTGATGATGATGCGAAGT 975
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 541 TNNAAATATTTTATATNNAAAAAATCATATATAAAGTTANTTGAATNTTATNATG 482
QY 976 AAAAGAAATCTCTAAAGAAATCTTACAATTA 1008
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 481 NAATTAATTCATATANAATAANANANTTA 449
RESULT 6
AI294319 684 bp mRNA EST 01-DEC-1998
LOCUS LP07659.5prime LP Drosophila melanogaster larval-early pupal POT2
DEFINITION Drosophila melanogaster cDNA clone LP07659 5prime similar to
U37541: Drosophila melanogaster complete mitochondrial genome, mRNA
sequence.
ACCESSION AI294319
NID 93943726
VERSION AI294319.1 GI:3943726
KEYWORDS EST:
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1. (bases 1 to 684)
AUTHORS Harvey, D., Hong, J., Evans-Holm, M., Pendleton, J., Su, C.,
Brokstein, P., Lewis, S. and Rubin, G.M.
BDGP/HMMI Drosophila EST Project
JOURNAL Unpublished (1997)
COMMENT On Jan 14, 1998 this sequence version replaced gi:1797824.
Contact: Harvey, D.
G. M. Rubin-Molecular and Cell Biology
University of California Berkeley
539 LSA, Berkeley, CA 94720-3200, USA
Fax: 510 643 9947
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
Plate: 76 row: E column: 11
High quality sequence stop: 623.
Location/Qualifiers
    1..684
FEATURES
    source

```


JOURNAL
MEDLINE
DNA Res. 5 (6), 335-340 (1998)
99156227
On Jan 14, 1998 this sequence version replaced q1:1797046.
COMMENT

PROJECT -	Discostereum discoideum CDAR project in Japan;
FEATURES	
Location/Qualifiers	
source	1..424

```

EASE COUNT
ORIGIN
256 a 24 c 13 g 130 t 1 others
/dev_stage="slug"
/done=140- directosystem discoideum 39 ("...")

```

Qy	685	TATTCAGAAAAATTTTCAAGGTGTTTAGTCAAAGATGATAAAGAATTGACTAAATGATTCT	744
Db	113	GAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATTTTAATCAGAAAAGGA	172

Qy 745 CATGTTAGTGATTATATTTATTAATAATTCGCGTACACCAATT 804
- - - - -
pb 173 TTTTTTAAATTAATAGAAATTAATAATAATTCCAAATCAATAATTAAT 232

QY 805 ACTGATTTTGAACATGAATTATTTTGTCGTGAATATAAATGAAAACCTGCATTTTAA 864
+
DB 233 ATCAATTAATAAAAAATTATAAAAAATTAATAATTAATAAAAAATAATAA-----AAAAAAT 287

[illegible][illegible]

Qy 985 A 985
—
408 7 408

RESULT 10
1156057

| LOCUS | AA550575 | 755 bp | DNA | EST | 11-AUG-1997 |
|------------|--|--------|-----|-----|-------------|
| DEFINITION | 1740m3 gmbpfbHB3.1, G. Roman Reddy Plasmodium falciparum genomic clone 1740m, mRNA sequence. | | | | |

NID 92320827

VERSION AA550575.1 GI:2320827
 KEYWORDS EST.
 SOURCE malaria parasite P. falciparum.
 ORGANISM Plasmodium falciparum
 REFERENCE 1 (bases 1 to 755)
 AUTHORS Eukaryota: Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 Dame, J.B., Annot, D.E., Bourke, P., Chakrabarti, D., Christodoulou, Z., Coppel, R., Cowman, A., Craig, A., Fischer, K., Foster, J., Goodman, N., Hinterberg, K., Holder, A.A., Holt, D., Kemp, D., Lanzer, M., Lim, A., Newbold, C., Ravetsch, J.V., Reddy, G.R., Rubio, J., Schuster, S.M., Su, X.-Z., Thompson, J.K., Vital, F., Wellem, T.E., and Werner, E.
 TITLE Current status of the Plasmodium falciparum genome project
 JOURNAL Mol. Biochem. Parasitol. 79, 1-12 (1996)
 MEDLINE 97001675
 COMMENT On May 18, 1995 this sequence version replaced gi:811125.

Contact: Dame JB
 Department of Pathobiology
 University of Florida
 2015 SW 16th Ave., Bldg. 1017, Gainesville, FL 32611-0880
 Tel: 352 392 4700
 Fax: 352 392 9704
 Email: dame@icbr.ifas.ufl.edu
 Seq primer: T3.

FEATURES
 source
 1..755
 /organism="Plasmodium falciparum"
 /db_xref="taxon:5833"
 /map="1"
 /clone="1740m"
 /lab_host="E. coli XLI-Blue"
 /note="Vector: pBluescript SK(+); Genomic DNA, from asynchronous blood stage parasites of the cloned Honduran HB3 isolate cultured in vitro, was digested with mung bean nuclease in the presence of 30% formamide at 50°C (Vernick, K.D., Imberski, R.B., and McCutchan, T.F. 1988. Nucleic Acids Research 16:6883-6896). The ends of the fragments were ligated using T4 DNA polymerase, and the dephosphorylated pBluescript SK(+). Recombinant plasmids transformed E. coli XLI-Blue."
 BASE COUNT 342 a 48 c 107 g 258 t
 ORIGIN

Query Match 5.1%; Score 51.8; DB 35; Length 755;
 Best Local Similarity 47.9%; Pred. No. 6.9;
 Matches 149; Conservative 0; Mismatches 162; Indels 0; Gaps 0;
 QY 597 TTATAAGCACCAGAAATGATCTGTGTATACTAATATGAATATGAATGAATGAAATGATTTG 656
 Db 160 TAATAAGAGACTTTATTTATCCATAATTAAGAAGGGCATATCTAATAATAATGA 219
 QY 657 GTCAATGGTATATTTGACTGGTTTATATTCAGAAATTTTCAAGTCTTTTACTCAA 716
 Db 220 ATCATATAAGGTATTCTATGATGAATATTTGTGAAGAAATGAAGAAACATATTA 279
 QY 717 AGATGATAAGAAATGACTAATGATTCATGTTAGTGTATTTATATTTATTAATCAAT 776
 Db 280 TAATATAGAAAAAGTATATAGAAATGTTGTAATAATATATATCAATGATTTAT 339
 QY 777 ATTTGAAATTTCCGTACACCCAAATTAATGATTTTGAAGATGAATATTTTGTGATGA 836
 Db 340 AAATAACATATTAATAAGCAAGAGCAATATTTATTACATATAGAACCTAGTGGTGT 399
 QY 837 ATATAATATCAAACTGTCATTTTAAAAAATTCATTTTACAAAAATATCCCTAGAAAA 896
 Db 400 ATTATATAAAATCATGAGAAATATTTTAAATAGCTTTTACGTTATGATCCAGAAAA 459
 QY 897 TTGGGATATTA 907
 Db 460 TGAAGATAATA 470

RESULT 11
 AU056862
 LOCUS
 DEFINITION
 AU056862 Oryza sativa mature leaf Nipponbare Oryza sativa CDNA
 clone S20934_1A, mRNA sequence.
 AU056862
 NID
 AU056862
 VERSION
 AU056862.1 GI:4715746
 KEYWORDS
 SOURCE EST.
 ORGANISM
 Oryza sativa.
 Oryza sativa

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Contact: Takuji Sasaki
 National Institute of Agrobiological Resources
 Rice Genome Research Program
 2-1-2 Kannondai, Tsukuba
 Ibaraki,
 Japan 305
 Tel: 0298-38-7441
 Fax: 0298-38-7468
 Email: tsasaki@abr.affrc.go.jp
 PROJECT "RGP".

FEATURES
 source
 1..712
 /organism="Oryza sativa"
 /strain="Nipponbare"
 /db_xref="taxon:4530"
 /clone="S20934_1A"
 /clone_lib="Oryza sativa mature leaf Nipponbare"
 /tissue_type="mature leaf"
 BASE COUNT 195 a 149 c 166 g 195 t 7 others
 ORIGIN

Query Match 5.1%; Score 51.8; DB 49; Length 712;
 Best Local Similarity 51.0%; Pred. No. 7.1; 138; Indels 3; Gaps 1;
 Matches 147; Conservative 0; Mismatches 138; Indels 3; Gaps 1;
 QY 408 GAGTTCAGACTTGAAATTTATTCATTCACAAGGGATATTCATCGTGATATAAAACCCAG 467
 Db 71 GATTTATGTCCTNAACTATCTCCATAGTAAGTGGATTATCCATCGAGATCTCAGCCTTC 130
 QY 468 TAATATTTTCTTTGCCCGGATGATATAACACACCGGATTTATGGAGATTTTGATTTG 527
 Db 131 TAATATCTGTCATGGGAGAGAGAGAAACATGGAAATTAAGAGATTGCTGATTTGG 190
 QY 528 TTATGATTTAAACTGCCACCTAAAGATGAACCCCTATGCCAAATATATGATGATC 587
 Db 191 ACTCGCTAGGATATATCAAGCTCCATTAAGCCATTAAGTATTAACGGGTGTTGT--- 248
 QY 588 TACAGGATTTTATAAGCACCAGAAATGATTCCTTGGTATAACTAATTAATTAATGAAT 647
 Db 248 TACCAITCGTATGGGCTCCAGAGTGTACTTGGGGCAAGCAGCTACACAAGTGTGT 307
 QY 648 TGATATTTGGTCATTTGGTATAATTTTCACTGGTTTATATTCAGAAA 695
 Db 308 TGATATGGGCACTTGGTGCATTTTGTGTAATTTGCTTACACTCAA 355

RESULT 12
 AU038989
 LOCUS
 DEFINITION

AU038989 579 bp mRNA EST 29-MAR-1999
 AU038989 Dictyostellium discoideum SS (H.Urushihara) Dictyostellium

ACCESSION AU001063
 NID 94157307
 VERSION AU001063.1 GI:4157307
 KEYWORDS EST.
 SOURCE domestic silkworm.
 ORGANISM Bombyx mori
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;
 Ditrysia; Bombycoidea; Bombycidae; Bombyx.
 REFERENCE 1 (bases 1 to 688)
 AUTHORS Mita, K., Morimyo, M., Shimada, T., Okano, K. and Maeda, S.
 TITLE Establishment of cDNA database of Bombyx mori
 JOURNAL Unpublished (1999)
 COMMENT On Oct 8, 1998 this sequence version replaced gi:3726239.

```

Contact: Mita K
Genome Research Group
National Institute of Radiological Sciences
Anagawa 4-9-1, Inage, Chiba 263-8555, Japan
Email: km.taue@nirs.go.jp
PROJECT = 'CREST project by JST'.
Location/Qualifiers
1..688
/organism="Bombyx mori"
/strain="p50(Daizo)"
/db_xref="taxon:7091"
/map="8"
/clone="fbf11331"
/clone_lib="Bombyx mori p50(Daizo)"
285 a 33 c 55 g 315 t
FEATURES
Source
BASE COUNT
ORIGIN

```

| Query Match | 5.08; | Score 51.4; | DB 45; | Length 688; |
|-----------------------|-----------------|---|-----------|-------------|
| Best Local Similarity | 48.3%; | Pred. No. 8.2; | | |
| Matches 208; | Conservative 0; | Mismatches 216; | Indels 7; | Gaps 2; |
| QY | 515 | ATTTTGATATTGTTATGATTTAAACTGGCACCTAAAGATGAACCCCTATGGCGAAT | 574 | |
| Db | 86 | ATTTTAAAGITTCGCAATTTAAAGAGTTAATTAATTAATAAGTTTATGTAGTAT | 145 | |
| QY | 575 | ATATTGATGATCTACAGGTTATTATTAAGCACAGAAATGATCTTGGTATAACTAAAT | 634 | |
| Db | 146 | A-AAATATTTAAATAAATTAATTTGAAATGAAATGTTATTCGTTTTTAAATATATCTAGTT | 204 | |
| QY | 635 | ATGAATATGAAATGATATTGGTCATTGGGTATATTTTGGCTGCTTTATATTCAGAA | 694 | |
| Db | 205 | TTTATAGAAAAAATTTAAATTTTATATATTTAGATTTTTTTTAAATTAATAATAATAA | 264 | |
| QY | 695 | ATTTTCAAAGTGTTTTATGCTCAAGATGATTAAGAAATGACTAATGATTCATGTTAGTG | 754 | |
| Db | 265 | AT-----ATATTTTAAATAAATAATTTAAGGATAAATCTTTAAATTTAAATTTTATAA | 318 | |
| QY | 755 | ATTTATATTTTAAATCAAAATATTTTCAAATTTTCGGTACACCCCAATTTAACTGATTTTG | 814 | |
| Db | 319 | TAATTTTAAATTAAGAAAGAAATTTATATAATTTATATTTGTTAATAAAATTTTATTTTATAT | 378 | |
| QY | 815 | AGATGAATATTTTGTGATGAATATAATAATGAAACCTTGCATTTTAAAAAATTCAAAT | 874 | |
| Db | 379 | AAATAAATTTTAACTAATAGAAAAATTTAAATTTTAAATTTTTTATATAAAAAATAAATTT | 438 | |
| QY | 875 | TACAAAAAATATCCTAGAAAAAGATTGGGATATTTATTTTAACTCGATGCAATGATGATTTAA | 934 | |
| Db | 439 | TAATTTAATAAATTAATTTAATGATGATAAACTTAGTAAAAATTTTAAATTAATAAAAAATTA | 498 | |
| QY | 935 | TGAAGAAATTT | 945 | |
| Db | 499 | TTAATTTTATTT | 509 | |

| | | | | |
|----------|----------|--------|------|-------------|
| RESULT | 15 | | | |
| AU002045 | | | | |
| LOCUS | AU002045 | 593 bp | mRNA | EST |
| | | | | 15-JAN-1999 |

| | |
|------------|---|
| DEFINITION | AU002045 Bombyx mori p50(Daizo) Bombyx mori cDNA clone mg0367, mRNA sequence. |
| ACCESSION | AU002045 |
| NID | g4158438 |
| VERSION | AU002045.1 GI:4158438 |
| KEYWORDS | EST. |
| SOURCE | domestic silkworm. |
| ORGANISM | Bombyx mori |
| | Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea; Bombycidae; Bombyx. |
| REFERENCE | 1 (bases 1 to 593) |
| AUTHORS | Mita, K., Morimyo, M., Shimada, T., Okano, K. and Maeda, S. |
| TITLE | Establishment of cDNA database of Bombyx mori |
| JOURNAL | Unpublished (1999) |
| COMMENT | On Oct 30, 1998 this sequence version replaced qi:3816581. |

Contact: Mita K.
Genome Research Group
National Institute of Radiological Sciences
Anagawa 4-9-1, Inage, Chiba 263-8555, Japan
Email: kmita@nirs.go.jp
PROJECT = 'CREST project by JST'.

| FEATURES | SOURCE |
|-------------------------|------------|
| 1. Feature 1 | Source 1 |
| 2. Feature 2 | Source 2 |
| 3. Feature 3 | Source 3 |
| 4. Feature 4 | Source 4 |
| 5. Feature 5 | Source 5 |
| 6. Feature 6 | Source 6 |
| 7. Feature 7 | Source 7 |
| 8. Feature 8 | Source 8 |
| 9. Feature 9 | Source 9 |
| 10. Feature 10 | Source 10 |
| 11. Feature 11 | Source 11 |
| 12. Feature 12 | Source 12 |
| 13. Feature 13 | Source 13 |
| 14. Feature 14 | Source 14 |
| 15. Feature 15 | Source 15 |
| 16. Feature 16 | Source 16 |
| 17. Feature 17 | Source 17 |
| 18. Feature 18 | Source 18 |
| 19. Feature 19 | Source 19 |
| 20. Feature 20 | Source 20 |
| 21. Feature 21 | Source 21 |
| 22. Feature 22 | Source 22 |
| 23. Feature 23 | Source 23 |
| 24. Feature 24 | Source 24 |
| 25. Feature 25 | Source 25 |
| 26. Feature 26 | Source 26 |
| 27. Feature 27 | Source 27 |
| 28. Feature 28 | Source 28 |
| 29. Feature 29 | Source 29 |
| 30. Feature 30 | Source 30 |
| 31. Feature 31 | Source 31 |
| 32. Feature 32 | Source 32 |
| 33. Feature 33 | Source 33 |
| 34. Feature 34 | Source 34 |
| 35. Feature 35 | Source 35 |
| 36. Feature 36 | Source 36 |
| 37. Feature 37 | Source 37 |
| 38. Feature 38 | Source 38 |
| 39. Feature 39 | Source 39 |
| 40. Feature 40 | Source 40 |
| 41. Feature 41 | Source 41 |
| 42. Feature 42 | Source 42 |
| 43. Feature 43 | Source 43 |
| 44. Feature 44 | Source 44 |
| 45. Feature 45 | Source 45 |
| 46. Feature 46 | Source 46 |
| 47. Feature 47 | Source 47 |
| 48. Feature 48 | Source 48 |
| 49. Feature 49 | Source 49 |
| 50. Feature 50 | Source 50 |
| 51. Feature 51 | Source 51 |
| 52. Feature 52 | Source 52 |
| 53. Feature 53 | Source 53 |
| 54. Feature 54 | Source 54 |
| 55. Feature 55 | Source 55 |
| 56. Feature 56 | Source 56 |
| 57. Feature 57 | Source 57 |
| 58. Feature 58 | Source 58 |
| 59. Feature 59 | Source 59 |
| 60. Feature 60 | Source 60 |
| 61. Feature 61 | Source 61 |
| 62. Feature 62 | Source 62 |
| 63. Feature 63 | Source 63 |
| 64. Feature 64 | Source 64 |
| 65. Feature 65 | Source 65 |
| 66. Feature 66 | Source 66 |
| 67. Feature 67 | Source 67 |
| 68. Feature 68 | Source 68 |
| 69. Feature 69 | Source 69 |
| 70. Feature 70 | Source 70 |
| 71. Feature 71 | Source 71 |
| 72. Feature 72 | Source 72 |
| 73. Feature 73 | Source 73 |
| 74. Feature 74 | Source 74 |
| 75. Feature 75 | Source 75 |
| 76. Feature 76 | Source 76 |
| 77. Feature 77 | Source 77 |
| 78. Feature 78 | Source 78 |
| 79. Feature 79 | Source 79 |
| 80. Feature 80 | Source 80 |
| 81. Feature 81 | Source 81 |
| 82. Feature 82 | Source 82 |
| 83. Feature 83 | Source 83 |
| 84. Feature 84 | Source 84 |
| 85. Feature 85 | Source 85 |
| 86. Feature 86 | Source 86 |
| 87. Feature 87 | Source 87 |
| 88. Feature 88 | Source 88 |
| 89. Feature 89 | Source 89 |
| 90. Feature 90 | Source 90 |
| 91. Feature 91 | Source 91 |
| 92. Feature 92 | Source 92 |
| 93. Feature 93 | Source 93 |
| 94. Feature 94 | Source 94 |
| 95. Feature 95 | Source 95 |
| 96. Feature 96 | Source 96 |
| 97. Feature 97 | Source 97 |
| 98. Feature 98 | Source 98 |
| 99. Feature 99 | Source 99 |
| 100. Feature 100 | Source 100 |

```

BASE COUNT      257 a 9 c 48 g 279 t
/organism="Bombyx mori"
/strain="p50(Daizo)"
/db xref="taxon:7091"
/clone="mg0367"
/clone_llb="Bombyx mori p50(Daizo)"

```

| Query Match | 5.0% | Score 51.4; | DB 45; | Length 593; |
|-----------------------|-----------------|---|-----------|-------------|
| Best Local Similarity | 48.3%; | Pred. No. 8.6; | | |
| Matches 208; | Conservative 0; | Mismatches 216; | Indels 7; | Gaps 2; |
| Qy | 515 | ATTTTGATATTGTTATGATTTAAACCTGCACCTAAAGATGAACCCCTATGGCGAAT | 574 | |
| Db | 157 | ATTTTAAAGTTTCTCGAATTTAAAGAGTTAAATTAATAAGTTATGTAGTAT | 216 | |
| Qy | 575 | ATATTGATGATACACAGGTATTTATAAAGCACCAGAAATGATTCTTGGTATATACTAAT | 634 | |
| Db | 217 | A-AATATTTTAAATAAATAATTTGAAATGAAATGTATTCTGTTTTTAAATATATCTAGTT | 275 | |
| Qy | 635 | ATGAATATGAATATGATTTTGGTCANTGGGTATAAATTTTCACGTGGTTATATATCAGAAA | 694 | |
| Db | 276 | TTTATAGAAAAAATAATTTTATTAATAGATTTTTTATTAAATTAATAATAA | 335 | |
| Qy | 695 | ATTTTCAAGTGTTTTAGTCAAAGATGATAAGAAATTGACTAATGATCTCATGTTAGTG | 754 | |
| Db | 336 | AT-----ATATTTTATAAATAAATTTAAGGATAATCTTTAAATTTAAATTTATAA | 389 | |
| Qy | 755 | ATTTATATTTATTAATCAATAATTTCAAAATTTCCGATACACCAAAATTTAACTGATTTG | 814 | |
| Db | 390 | TAATTTTATTAAGAGAAATTTATATAATTTATTTGTTAAAGAAATTTATTTATAT | 449 | |
| Qy | 815 | AAGATGAATTTATTTGATGAATATAATGAAGAACTTGCATTTTAAAAAAATTCAT | 874 | |
| Db | 450 | AAATAAATTTTACTAATAAGAAAAATTTAATTTATTTTATTTATATAAAATAAATTT | 509 | |
| Qy | 875 | TACAAAAATATCCTAGAAAAAGATTGGGATATATTTTACCTCGATGCAATGATATTA | 934 | |
| Db | 510 | TAATTTAAATTAATTAATTAATGATAAAATTAGTAAAAATTTATTAATTAATAAAAA | 569 | |
| Qy | 935 | TGAAAGAAATTT | 945 | |
| Db | 570 | TTAATTTATTT | 580 | |

Search completed: November 6, 1999, 07:35:44

Sat Nov 6 11:07:19 1999

us-09-072-994-l3.rst

Page 10

Job time: 19381 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 1999, 21:13:23 : Search time 23.99 Seconds
(without alignments)
4419.350 Million cell updates/sec

Title: US-09-072-994-13
Perfect score: 1019
Sequence: 1 ATGAAGTTCAGATTATTA.....TTACAAATTAATGTTGATTG 1019
Scoring table: IDENTITY_NUC
Searched: 192659 seqs, 52021692 residues
Database : Issued Patents NA.*
1: /cgn2.6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2.6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2.6/ptodata/1/ina/5C_COMB.seq.*
4: /cgn2.6/ptodata/1/ina/5D_COMB.seq.*
5: /cgn2.6/ptodata/1/ina/PTUS9_COMB.seq.*
6: /cgn2.6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------------------------|
| 1 | 56.2 | 5.5 | 8920 | 3 | US-08-446-855A-1 Sequence 1, Appl |
| 2 | 53 | 5.2 | 1070 | 3 | US-08-463-090B-1 Sequence 5, Appl |
| 3 | 51.4 | 5.0 | 2483 | 2 | US-08-526-964-1 Sequence 1, Appl |
| 4 | 51.4 | 5.0 | 2483 | 3 | US-08-946-611-1 Sequence 5, Appl |
| 5 | 50.4 | 4.9 | 8920 | 3 | US-08-446-855A-1 Sequence 1, Appl |
| 6 | 49 | 4.8 | 1511 | 1 | US-07-991-867B-8 Sequence 8, Appl |
| 7 | 49 | 4.8 | 1511 | 2 | US-08-107-755A-8 Sequence 8, Appl |
| 8 | 49 | 4.8 | 1511 | 4 | US-08-544-332-8 Sequence 1, Appl |
| 9 | 48.8 | 4.8 | 8457 | 1 | US-07-991-867B-1 Sequence 1, Appl |
| 10 | 48.8 | 4.8 | 1395 | 1 | US-07-991-867B-23 Sequence 25, Appl |
| 11 | 48.8 | 4.8 | 1395 | 2 | US-08-107-755A-1 Sequence 1, Appl |
| 12 | 48.8 | 4.8 | 1395 | 4 | US-08-107-755A-25 Sequence 25, Appl |
| 13 | 48.8 | 4.8 | 8457 | 4 | US-08-544-332-1 Sequence 9, Appl |
| 14 | 48.8 | 4.8 | 1395 | 4 | US-08-544-332-25 Sequence 9, Appl |
| 15 | 48 | 4.7 | 7101 | 2 | US-08-480-604A-9 Sequence 9, Appl |
| 16 | 47.4 | 4.7 | 7101 | 4 | US-07-867-106-2 Sequence 2, Appl |
| 17 | 47.4 | 4.7 | 5852 | 1 | US-07-867-106-2 Sequence 36, Appl |
| 18 | 46.8 | 4.6 | 319 | 1 | US-07-872-678A-36 Sequence 14, Appl |
| 19 | 46.8 | 4.6 | 319 | 1 | US-07-593-657-14 Sequence 1, Appl |
| 20 | 46.8 | 4.6 | 8457 | 1 | US-07-991-867B-1 Sequence 1, Appl |
| 21 | 46.8 | 4.6 | 1689 | 1 | US-07-991-867B-41 Sequence 41, Appl |
| 22 | 46.8 | 4.6 | 8457 | 4 | US-08-544-332-1 Sequence 1, Appl |
| 23 | 46.8 | 4.6 | 1689 | 4 | US-08-544-332-41 Sequence 41, Appl |
| 24 | 46.6 | 4.6 | 7351 | 2 | US-08-224-391-83 Sequence 83, Appl |
| 25 | 46.6 | 4.6 | 7351 | 3 | US-08-484-304-83 Sequence 83, Appl |
| 26 | 46.6 | 4.6 | 7351 | 3 | US-08-184-009-127 Sequence 127, Appl |
| 27 | 46.6 | 4.6 | 4637 | 3 | US-08-565-398-16 Sequence 16, Appl |
| 28 | 46.6 | 4.6 | 4637 | 3 | US-08-565-398-39 Sequence 39, Appl |
| 29 | 46.6 | 4.6 | 4231 | 3 | US-08-417-210A-80 Sequence 80, Appl |
| 30 | 46.6 | 4.6 | 7351 | 4 | US-08-458-256-127 Sequence 27, Appl |
| 31 | 46.2 | 4.5 | 3891 | 2 | US-08-480-604A-27 Sequence 27, Appl |
| 32 | 46.2 | 4.5 | 3891 | 1 | US-07-867-106-2 Sequence 2, Appl |
| 33 | 46 | 4.5 | 5852 | 1 | US-07-867-106-2 Sequence 32, Appl |
| 34 | 45.6 | 4.5 | 660 | 2 | US-08-107-755A-32 Sequence 32, Appl |
| 35 | 45.6 | 4.5 | 660 | 4 | US-08-544-332-32 Sequence 32, Appl |
| 36 | 45.6 | 4.5 | 660 | 2 | US-08-544-332-32 Sequence 32, Appl |
| 37 | 45.4 | 4.5 | 9636 | 2 | US-08-333-170B-1 Sequence 1, Appl |

| | | | | | |
|------|------|-----|------|---|-----------------------------------|
| C 38 | 44.8 | 4.4 | 1511 | 1 | US-07-991-867B-8 Sequence 8, Appl |
| C 39 | 44.8 | 4.4 | 1511 | 2 | US-08-107-755A-8 Sequence 8, Appl |
| C 40 | 44.8 | 4.4 | 4098 | 4 | US-08-603-106-4 Sequence 4, Appl |
| C 41 | 44.8 | 4.4 | 1511 | 4 | US-08-544-332-8 Sequence 8, Appl |
| C 42 | 44.6 | 4.4 | 2448 | 2 | US-08-526-964-2 Sequence 2, Appl |
| C 43 | 44.6 | 4.4 | 5495 | 2 | US-08-602-010A-1 Sequence 1, Appl |
| C 44 | 44.6 | 4.4 | 5495 | 2 | US-08-602-010A-2 Sequence 2, Appl |
| C 45 | 44.6 | 4.4 | 1203 | 2 | US-08-602-010A-7 Sequence 7, Appl |

ALIGNMENTS

RESULT 1
US-08-446-855A-1
Sequence 1, Application US/08446855A
Patent No. 5895711
GENERAL INFORMATION:
APPLICANT: Stewart, Thomas S
APPLICANT: O'Sullivan, William J
TITLE OF INVENTION: Nucleotide sequence encoding carbamoyl phosphate synthetase II
TITLE OF INVENTION: Nucleotide sequence encoding carbamoyl phosphate synthetase II
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon & Vanderhye PC
STREET: 1100 No. 5849573th Glebe Road, 8th Floor
CITY: Arlington
STATE: Virginia
COUNTRY: USA
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.24
CURRENT APPLICATION NUMBER: US/08/446,855A
FILING DATE: 06-Jul-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mitchard, Leonard C
REGISTRATION NUMBER: 29.009
REFERENCE/DOCKET NUMBER: 47-80
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 8920 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: genomic
US-08-446-855A-1

Query Match Score 56.2; DB 3; Length 8920;
Best Local Similarity 45.4; Pred. No. 0.0017;
Matches 328; Conservative 0; Mismatches 383; Indels 12; Gaps 3;

| | |
|---------|---|
| QY 153 | AGAAATTTTACTTAAACCTTTGAACACACATCCAAACATATTTGAATATTTAATGA 212' |
| Db 5854 | AGGGAATTAATAAATATGGATATATCTAAGGAAAGAAAGATATCTAATAAAAA 5913 |
| QY 213 | TCTTAAATTTATGATGATGATATATAGTACCAAAATTTGATCGTATGATTTGATGA 272 |
| Db 5914 | CCCTGGTAAATCTTATATGTTGATGATCCCTGATATTAATAATGAACAAATTAATA 5973 |
| QY 273 | ATTCATGAATTAACAAATTTGTAACGAACACACATATTTATTTATGATTTAATG 332 |
| Db 5974 | GATGAAGAGTTAATAGATAACGAAATTTAATATGATGAATATAATAATATGTTAATAT 6033 |

QY 333 TAATCTTGTAGTAATCAATATACACTTGTCTAATGAAATG-----AAGAAAAAGAT 384
DB 6034 GAATTTCTTCTAATATATATCTAGTCTGTAATGGAAGAGGATAGAAATGATAA 6093
QY 385 ATCAAAATTAAGTAAATCAATGAGTTCAGGA-CTTGAATTTATTCATTCACAGGAT 443
DB 6094 TTTAGAAATGATGTTATGAAATAATATGATGATACATACAAACATATATATCTGTT 6153
QY 444 AATTCATCTGATATAAACCAGTAATATTTCTTTGCCCGGATGATATAACACAAAC 503
DB 6154 AACATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6213
QY 504 GATATCTGAGATTTTGAATTTTGTATGATTTTAACTCCACCTTAAGATGAACCC 563
DB 6214 CATCATGAGAGGCCATGAGAAATGTTTAAATTAATAAATAAATAAATAAATAAATAA 6270
QY 564 TATGCGGAATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 623
DB 6271 TATGGAATAATTAAGGAAATAATTAACACAAAGATATCATATGTTATCATATAA 6330
QY 624 TATACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 633
DB 6331 AAGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6390
QY 684 ATATTCAGAAATTTTCAGATGTTTTCAGATGATGATGATGATGATGATGATGATG 743
DB 6391 TATATAATAAAGAAATATTAAGATGATGATGATGATGATGATGATGATGATGATG 6450
QY 744 TCATGTTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 803
DB 6451 ACAAGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6510
QY 804 AACTGATTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 863
DB 6511 TGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6570
QY 864 AAA 866
DB 6571 TAA 6573

RESULT 2
US-08-453-090B-5
; Sequence 5, Application US/08463090B
; Patent No. 5801015
; GENERAL INFORMATION:
; APPLICANT: Cottarel, Guillaume
; APPLICANT: Damagnez, Veronique
; APPLICANT: Draetta, Guilio
; TITLE OF INVENTION: Cell-Cycle Regulatory Proteins from
; TITLE OF INVENTION: Human Pathogens, and Uses Related Thereto
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley, Hoag & Elliot, LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,090B
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: G06F03/00
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent Matthew P.
; REGISTRATION NUMBER: 36 709
; REFERENCE/DOCKET NUMBER: MIV032.01
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 832-1299
TELEFAX: (617) 832-7000
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1070 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
MOLECULE TYPE: CDNA
FEATURES:
NAME/KEY: CDS
LOCATION: 30..1058
US-08-463-090B-5
Query Match 5.28; Score 53; DB 3; Length 1070;
Best Local Similarity 53.38; Pred. No. 0.0055;
Matches 172; Conservative 0; Mismatches 130; Indels 21; Gaps 2;
QY 362 CTAATGAATTTGAAGAAAGATATCAATTAATGTTAAATCAATGATTCAGGCTTG 421
DB 373 CGATGTTTCAATCAGCAGATATCAATCAATGCTTTTAAATCAGATTCAGTGGATAC 432
QY 422 AATTTATTCATTCACAGGATTAATTCATCGTATATAAAACCCAGTAATATTTCTTTG 481
DB 433 ATCATGTCATCGAATTTTATTTTACATCGTATTTGAACCAAAATATTTATTATTGG 492
QY 482 CCGGGGATGATATAACACACCGGATTTTGGAGATTTTGTATTTGTTATGATTTAAAC 541
DB 493 CACCGGATG-----ACAATGAAATAGCGGATTTTGGCTTCACGAGCTTTGGTAA 546
QY 542 TGCCACCTTAAAGATGAACCCCTTATGCGGAAATATATTGATGATCTACAGGTATTTATA 601
DB 547 ATCTTAATGAAGATTTATCATCTA-----ATGTTGTCACATAGATGATATA 591
QY 602 AAGCACCAGATGATCTCTGTTGATATAACTAATATGATGATGATGATGATGATGATG 661
DB 592 GAGCCCTGAAATATTATTGTTGTCGACATACATACATGAGCAGTTGATATCTGTC 651
QY 662 TGGGTATAATTTGACTGTTTA 684
DB 652 TAGGTATAATTTTCTGTAATA 674

RESULT 3
US-08-526-964-1
; Sequence 1, Application US/08526964
; Patent No. 5698421
; GENERAL INFORMATION:
; APPLICANT: Lambowitz, Alan M
; APPLICANT: Zimmerly, Steven
; APPLICANT: Guo, Huatao
; APPLICANT: Yang, Jian
; TITLE OF INVENTION: Nucleotide Integrase Preparation
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calfee, Halter & Griswold
; STREET: 800 Superior Avenue
; CITY: Cleveland
; STATE: Ohio
; COUNTRY: USA
; ZIP: 44114
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/526,964
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Goltick, Mary E

Query Match 5.0%; Score 51.4; DB 2; Length 2483;
Best Local Similarity 45.5%; Pred. No. 0.013;
Matches 338; Conservative 0; Mismatches 396; Indels 9

5

| | | | | |
|----|------|----------|---|------|
| QY | 281 | AAATTC | AAAAATATTTGTAACAGCAACACACAGATTTATTTATGGTATTAATAGGTAATCTTG | 340 |
| Db | | | | |
| QY | 892 | AAATTTG | CAAGAAGATGATGAGAAATAATATAGAAATATCTATAATATAGTTCTCTTT | 951 |
| Db | | | | |
| QY | 341 | TTAGTAAT | CAATATACACTTGCCTAAATGAATGAAATGGAAGAAAGATATCAAAATATAGGTTAA | 400 |
| Db | | | | |
| QY | 952 | ATTATTC | TATCGATTTAGACCTTAACCTATCTCTTTTAACAGCATTTATTCAAATGTAAAA | 1011 |
| Db | | | | |
| QY | 401 | AATCAAT | GAGTTCAGAGCTTGAATTTATTTCAATTCACAAGGGATTAATTCATCGTGATATAA | 460 |
| Db | | | | |
| QY | 1012 | ATTATAT | GCAATACGTAAATGATTTATTTAAAGTAGATTTAAATAAATGCTTTGTGATACAA | 1071 |
| Db | | | | |
| QY | 461 | AAACCA | GTAATATTTCTTTGCGCGGATGATATACACACCGGATTAATTCGGAGTTTG | 520 |
| Db | | | | |
| QY | 1072 | TTCCAC | ATATATGTAATTT - AATGTATTAATGAGAGNATCAAGATATAAGGTTTCAT | 1129 |
| Db | | | | |
| QY | 521 | ATATTTG | TTGTAATTTAAACTGCCACTTAAGATGAACCCCTATGCGGAATATATATG | 580 |
| Db | | | | |
| QY | 1130 | AGACTT | ATATAT - ATAAATTTAAGAGCTGGGATATGTTGATAAAATTAATATATTCATA | 1188 |
| Db | | | | |
| QY | 581 | ATGTATC | TACAGTATTTATAAGCACACAGAAATGATCTTGGTATA --- ACTAATATATG | 637 |
| Db | | | | |
| QY | 1189 | ATACACT | TTTAGGAATCCCAAGGAGTGTGTGCAGTCTCTATTTATGTAAATATTTTTT | 1248 |
| Db | | | | |
| QY | 638 | AAATGA | AAATTCATTTTGGTCAATGGGTATAAATTTTGACTCGTGTATATATTCAGAAAT - | 697 |
| Db | | | | |
| QY | 1249 | TAGATA | AAATTTAGATAAATATTTAGAAAAATAATTTGAGATGAATTCAACTACTGGAATA | 1308 |
| Db | | | | |
| QY | 697 | - - -TTT | CAAAAGTCTTTTAGTCAAAAGATGATAAAGAAATTTGACTATGATCTCATGTATG | 754 |
| Db | | | | |
| QY | 1309 | TGCTAT | AGAGTGAAGAAATCCCAATTTATAGTTTATCATCTAAAAATTTATAGATGTA | 1368 |
| Db | | | | |
| QY | 755 | ATTTAT | ATATTTAAATCAAAATTTTGAAAATTTTCGGTACACCCCAATTTAACTGATTTTG | 814 |
| Db | | | | |
| QY | 1369 | AATTTAT | CTCAAAAATTTAAATTTAGATTAAAGAGACCAATACCCAAAGAAATATGG | 1428 |
| Db | | | | |
| QY | 815 | AAAGAT | GAATTTTGTGATGAATATATATATGAACACTTGCATTTTAAANAATTCAAAT | 874 |
| Db | | | | |
| QY | 1429 | GATCTG | ATAAAGTATTATTTTGTAGATATGCTGATGATATATATCATG | 1488 |
| Db | | | | |
| QY | 875 | TACAAA | ANATCTCTAGAAAAGATTTGGGATATATTTTACTCGATGCAATGATGATTAA | 934 |
| Db | | | | |
| QY | 1489 | GTGTAT | TGGGTCTCATATATGATTTGTAATAATTTTTTAACGATATTTAAACTCTCTTAA | 1548 |
| Db | | | | |
| QY | 935 | TGAAGA | ANATTTTTTACCAGATGATTTAGATATGATCGAAGTAAAAAGAAATAACTCTTAAG | 994 |
| Db | | | | |
| QY | 1549 | AAGAAA | ATTTTAGGTATGTCAATTAATATAGATAAATCCGTTATTAACCATCTTAAAGAG | 1608 |
| Db | | | | |
| QY | 995 | AAATCT | TACAATTAATGTGGAT | 1017 |
| Db | | | | |
| QY | 1609 | GAGTTAG | TTTTTTTAGGGTATGAT | 1631 |
| Db | | | | |

SEQUENCE CHARACTERISTICS:
LENGTH: 2483 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
HYPOTHETICAL: NO
US-08-946-617-1

Query Match 5.0%; Score 51.4; DB 3; Length 2483;
Best Local Similarity 45.5%; Pred. No. 0.013;
Matches 338; Conservative 0; Mismatches 396; Indels 9; Gaps 4;
QY 281 AAATTACAATAATGTAACGACACACAGATTTATTTATGGTATTAATGTAATCTTG 340
DB 892 AAATGTCAGAAAGTATGAGATATATTGAATATATCTATATATAGTTCTCTT 951
QY 341 TTAGTAATCAATATACACTTCTGTAAGAAATTGAAGAAAGATATCAAAATATGTTAA 400
DB 952 ATTATCTCATGGATTTAGACCTAATCTATCTGTTTAAACAGCTATTATCAATGTAATA 1011
QY 401 AATCAATGAGTTTCAGGACTTGAATTTATTCATTCACAGGGATTAATTCATCGTATATA 460
DB 1012 ATTATGCAATCTGTAATGATTTTAAAGTAGATTTAAATAGATTTAAATGCTTTGATACAA 1071
QY 461 AACCCAGTAATATTTCTTTCCCGGGATGATATAACACACCGGATTTATGGAGATTTTG 520
DB 1072 TTCCACATAATATGTTAAT--AATGTAATTAAGAGATCAAGATTAAGGTTTCAT 1129
QY 521 ATATGTTATGATTTAAACCTCCACCTAAAGATGAACCCCTATGGCGAATATATG 580
DB 1130 AGACTTATAT--ATAAATTTAAGAGCTGGATGTTGATATAAATAAATAATATCATA 1188
QY 581 ATGATCTACAGGATTTATAAGACACCAAGATTTGATCTTGGTATA--ACTAATATG 637
DB 1189 ATACACITTAGGAATTCCTCAGGTAGTGTGTCAGTCTATTTTATGTAATATTTT 1248
QY 638 AATATCAATATGATTTGTCATCTGGGTATATTTTGCCTGTTTATATTCAGAAAT- 697
DB 1249 TAGATATATGATTAATATTTAGAAATAATTTGAGATGATTTCAATCTGGAATA 1308
QY 697 --TTTCAAGTGTTTAGTCAAGATGATAAGAAATTCAGTAATCTCATGTTAGTG 754
DB 1309 TGCTAATAGAGGTAGAAATCCCAATTTATAATAGTTTATCATCTAATAATTTATAGATGA 1369
QY 755 ATTATATTTTAAATCAATATTTGAAATTTCCGTACACCAATTTAACTGATTTG 814
DB 1369 AATATATCTGAAATTTAAATTTGATTAGATTAAGAGACCATTTCCCAAGAAATATGG 1428
QY 815 AAGATGAATTTATTTGTGATGATATAATAATGAAGAACTTGCATTTTAAAAAATTCAT 874
DB 1429 GATCTGATAAAGTTTAAAGAGCTTATTTGTTAGATATGCTGATCATATTTATCAT 1488
QY 875 TACAAAAATATCTAGAAAGATTTGGATATTTATTTTACCTCGATGCAATGATTTAA 934
DB 1489 GTGTAATGGTCTCAATGATTTGTAATAATTTTAAACGATATTAATAACTTCTTAA 1548
QY 935 TGAAGAGATTTTACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 994
DB 1549 AAGAAATTTAGGTATGCTAATTAATAGATAAATCCGTTTATTAACATTTCTAAGAG 1608
QY 995 AATCTTACAATTAATGTTGAT 1017
DB 1609 GAGTTAGTTTTTATGGGTATGAT 1631

RESULT 5
US-08-446-855A-1/c
Sequence 1, Application US/08446855A
Patent No. 5849573
GENERAL INFORMATION:
APPLICANT: Stewart, Thomas S

APPLICANT: Flores, Maria V
APPLICANT: O'Sullivan, William J
TITLE OF INVENTION: Nucleotide sequence encoding carbamoyl
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon & Vanderhye PC
STREET: 1100 No. 5849573th Glebe Road, 8th Floor
CITY: Arlington
STATE: Virginia
COUNTRY: USA
ZIP: 22201-4714

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.24
CURRENT APPLICATION DATA: US/08/446,855A
APPLICATION NUMBER: US/08/446,855A
FILING DATE: 06-Jul-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mitchard, Leonard C
REGISTRATION NUMBER: 29,009
REFERENCE/DOCKET NUMBER: 47-80
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 8920 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: genomic
US-08-446-855A-1

Query Match 4.9%; Score 50.4; DB 3; Length 8920;
Best Local Similarity 48.0%; Pred. No. 0.024;
Matches 144; Conservative 0; Mismatches 156; Indels 0; Gaps 0;
QY 102 TCTTAAATAGTTGATGAAGATTTTCAGTCTTCCACACATTCATCCATCGAGAAATTT 161
DB 8924 TTTTAAATAATCATCAATTTATAAATTTATATAATTAAGAATTTATPACATTT 8765
QY 162 TATACTTAAATCTTGAACACATCCAAACATPAATTTGAATTTTAAATGATCTTAAAT 221
DB 8764 AATGTTTATATTTTATTTTATATAATTAATAATTTTAAATTAATAAATTAAT 8705
QY 222 TTATGATGATTTATATTTAGTCACCAATTTGATCGTTATGATTTGAGTCAATTTGATGA 281
DB 8704 TTGTAATTTGAAAAACCATTTTGGTTTATACATATGAGTGAATAATAAATTTGTATA 8645
QY 282 AATTAACAATAATTTGTAACGAACACACCATTTTATTTATGTTATTAATGTTATCTTGT 341
DB 8644 TATAACAATAATTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 8585
QY 342 TAGTATCAATATACATTTGCTATGTAATGAAATTTGAAGAAAGATATCAAAATTTATGTTAA 401
DB 8584 CATTTATTAATTTAAACGTAATAATAATAATAATAATAATAATAATAATAATAATA 8525

RESULT 6
US-07-991-867B-8
Sequence 8, Application US/07991867B
Patent No. 5476781
GENERAL INFORMATION:
APPLICANT: Moyer, Richard W.
APPLICANT: Hall, Richard L.
APPLICANT: Guhl, Michael E.
TITLE OF INVENTION: NO. 5476781el Entomopoxvirus Expression System
NUMBER OF SEQUENCES: 66

TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1511 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE: Amsacta moorei entomopoxvirus
FEATURE:
NAME/KEY: CDS
LOCATION: complement (18..218)
FEATURE:
NAME/KEY: CDS
LOCATION: complement (234..782)
FEATURE:
NAME/KEY: CDS
LOCATION: 852..1511
US-08-107-755A-8

Query Match 4.9%; Score 49; DB 2; Length 1511;
Best Local Similarity 43.4%; Pred. No. 0.035;
Matches 325; Conservative 0; Mismatches 420; Indels 4; Gaps 2;
QY 139 CATTCAATCCATCGAGAAATTTTATATCTTAAACTTTGAAACCCATCCCAAAATATT 198
DB 761 CAATTATTAATCTTCGCGATTTTTTTTATTTATTTGATATATTTTTCAAAAAAAT 820
QY 199 GAATATTTTAAATGCTTAAATTTATGATGATGATATATATGATCACCAGATGTCGT 258
DB 821 TAATCAATGAAAAAATAAATAATATCAAAATGATTTTAAATTTCTGATATAATTT 880
QY 259 TATGATTTGAGTCAATGATGAAATATCAAAATATGTAACGAAACACAGATTTATT 318
DB 881 AATAAATATTTAAATATTTATATTTTAAAAAATAATATAACAGAGATAATGTTAT 940
QY 319 TATGATTTTAAATGATCTTTGATGATCAATATATACACTGCG---TAATGAATTTGAA 375
DB 941 TAATATTAATATTTAAAAAATTTAGTTAATTTAGAAAGATTTGATATATATATATGA 1000
QY 376 GAAAGATATCAATATGTTGTTAAATCAATGATTCAGGACTTGAATTTATTCATCA 435
DB 1001 TATATATTTTAAATATATCCAGAAATATTAAGTTTATATATTTCAATTTAA 1060
QY 436 CAAGGATATTCATGATATATTAACCCAGTATATTTTCTTCCCGGAGATATATA 495
DB 1061 TATATTAATTTAAATTTTATACAAAAATTAACAAATATTAAGTTTATATATTTCA 1120
QY 496 ACAC-AACCGATTTATGGAGATTTGATATTTGATGATTTAAACTGCCACCTAAGA 554
DB 1121 TAACAAAAATAGCAATATAAGTATATATATATACCATCTTCTATAGAAATTTTAAAT 1180
QY 555 TGAACCCCTATGCGGAATATATGATGATATCTACAGGATTTTATTAAGACACCAAGAT 614
DB 1181 TGAATCATGTAATAAATGACTATATATTTTAAATTAATTAAGTAAATTTAAAAAAT 1240
QY 615 GATTCCTGTTAACTAATATGATATGAAATGATATTTGCTCATTTGGGTATAATTT 674
DB 1241 AATAATATCTAAAAATAAATTTGGTAACCTTAAATAAGTTTTCCTTATAGTATAGTTGA 1300
QY 675 GACTGGTTTATATTCAGAAATTTTCAAGTGTTTTATGTCAGAAAGATGATTAAGAAAT 734
DB 1301 GTTAAATATGCAATCAATATAAAGATTTAATAATTTATAGAAAAATTAATTAATTT 1360
QY 735 TAATGATCTCATGTTAGTATTTATTTATTAATCAATATTTGAAATTTTCGGTAC 794
DB 1361 AAAAAAATAGATATATCTTCAATGTTAAAAAATAATATATACATTTGATAAAAATTC 1420
QY 795 ACCCAATTTACTGATTTTGAAGATGAATTTATTTGTTGATGAATATATATGAAGAACTT 854

DB 1421 AAAAGTATAAATCATTTATGTTATCATCATATATAAGAAATTAATTAATTAATA 1480
QY 855 GCATTTTAAAAAATTCATTTACAAAAT 883
DB 1481 AAATTTATCAATATATATGAATATGAAT 1509
RESULT
US-08-544-332-8
Sequence 8, Application US/08544332
Patent No. 5935777
GENERAL INFORMATION:
APPLICANT: Moyer, Richard W.
APPLICANT: Hall, Richard L.
APPLICANT: Gruidl, Michael E.
TITLE OF INVENTION: NO. 5935777el Entomopoxvirus Expression System
NUMBER OF SEQUENCES: 77
CORRESPONDENCE ADDRESS:
ADDRESSEE: Gerard H. Bencen
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/544,332
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/991,867
FILING DATE: 07-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/107,755
FILING DATE: 19-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO 92/14818
FILING DATE: 12-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,685
FILING DATE: 30-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/657,584
ATTORNEY/AGENT INFORMATION:
NAME: Bencen, Gerard H.
REGISTRATION NUMBER: 35,746
REFERENCE/DOCKET NUMBER: UFI14.C4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1511 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Amsacta moorei entomopoxvirus
FEATURE:
NAME/KEY: CDS
LOCATION: complement (18..218)
FEATURE:
NAME/KEY: CDS
LOCATION: complement (234..782)
FEATURE:
NAME/KEY: CDS

NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UF114.C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6768 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE: Amsacta moorei entomopoxvirus
FEATURE:
NAME/KEY: CDS
LOCATION: complement (65...1459)
FEATURE:
NAME/KEY: CDS
LOCATION: 1474..2151
FEATURE:
NAME/KEY: CDS
LOCATION: complement (2239..2475)
FEATURE:
NAME/KEY: CDS
LOCATION: 2502..2987
FEATURE:
NAME/KEY: CDS
LOCATION: 3080..6091
FEATURE:
NAME/KEY: CDS
LOCATION: complement (6277..6768)
US-08-107-755A-1

Query Match 4.8%; Score 48.8; DB 2; Length 6768;
Best Local Similarity 47.7%; Pred. No. 0.048;
Matches 207; Conservative 0; Mismatches 222; Indels 5; Gaps 2;
Qy 475 TTCCTTCCCGGGATGATATACACACCCGATTTGGAGATTTTGATTTGTTATGAT 534
Db 614 TTATATCCAGACTTAAACAAATAAATAATATGTTTATGATCTGATGAC 555
Qy 535 TTAACACTGCCACTTAAGATGAACCCCTATGGCGAATATATGATGATCTACAGGT 594
Db 554 TATAAAGCCA--TAACACTACATAGTACTCTCTATACTAATTAICTTATTATAC 497
Qy 595 ATTTAAGCACCAGATGATCTTGTATACATTAATGATGATGAATGATGATTT 654
Db 496 ATTGATGATCTATCAGATATTAATGATATATTAAGATGTAATTAATTAATTTT 437
Qy 655 TGGTCATTTGGGTATAATTTTGACTGGTTTATATTCAGAAATTTTCAAAAGTGTAGTC 714
Db 436 AATACCTTTGAATTAACAATATGATAGTCCCGATTTGGTATGTTAATAATATTTA 377
Qy 715 AAAGATGATAAAGAAATGATGATTCATGTTAGTATGATTTATATTTAATAACAA 774
Db 376 TATTATATAGTTTATTTCAACATAAAATCTAAATTTGAATTTAAAAAATATATTTCT 317
Qy 775 ATATTGAAATTTCCGTACACCCAAATTTAACTGATTTTGAAGATGAATTTTGGAT 834
Db 316 ATGAGTTTATTTGATTTATTTGGCT--AGTAGTTATAGAGTGCATTTATTCT 260
Qy 835 GAATATATATGAATGAAATTCATTTTAAAAAATTCATTTACAAAAATCTCTAGAAA 894
Db 259 AGATATGATATTAATAGTATTGACGAATTTAAAAATACATTTAGAAATATTTAATAAAA 200
Qy 895 GATTGGGATATTAT 908
Db 199 AATAAAGAGTTAT 186

RESULT 12
US-08-107-755A-25/C
Sequence 25, Application US/08107755A
Patent No. 5721352
GENERAL INFORMATION:
APPLICANT: Moyer, Richard W.
APPLICANT: Hall, Richard L.
APPLICANT: Gruidl, Michael E.
TITLE OF INVENTION: No. 5721352a1 Entomopoxvirus Expression System
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: U.S.A.
ZIP: 32606
COMPUTER READABLE FORM:
MOLECULE TYPE: IBM copy disk
COMPUTER: IBM compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/107,755A
FILING DATE: 19-AUG-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,658
FILING DATE: 30-JAN-1992
APPLICATION DATA:
APPLICATION NUMBER: US 07/657,584
FILING DATE: 19-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UF114.C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 1395 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
MOLECULE TYPE: unknown
US-08-107-755A-25

Query Match 4.8%; Score 48.8; DB 2; Length 1395;
Best Local Similarity 47.7%; Pred. No. 0.039;
Matches 207; Conservative 0; Mismatches 222; Indels 5; Gaps 2;
Qy 475 TTCCTTCCCGGGATGATATACACACCCGATTTGGAGATTTTGATTTGTTATGAT 534
Db 550 TTATATCCAGACTTAAACAAATAAATAATATGTTTATGATCTGATGAC 491
Qy 535 TTAACACTGCCACTTAAGATGAACCCCTATGGCGAATATATGATGATCTACAGGT 594
Db 490 TATAAAGCCA--TAACACTACATAGTACTCTCTATACTAATTAICTTATTATAC 433
Qy 595 ATTTAAGCACCAGATGATTTGTTGTTATACATTAATGATGATGAATGATGATTT 654
Db 432 ATTGATGATCTATCAGATATTTAAATGATATATTTAAAAATGTAATTTATTCT 373
Qy 655 TGGTCATTTGGGTATAATTTTGACTGGTTTATATTCAGAAATTTTCAAAAGTGTAGTC 714
Db 372 AATACCTTTGAATTAACAATATGATAGTCCCGATTTGGTATGTTAATAATAATTTT 313
Qy 715 AAAGATGATAAAGAAATGATGATTCATGTTAGTATGATTTATTTAATAACAA 774
Db 312 TATTATAGTTTATTTCAACATAAATCTAAATTTGAATTTAAAAAATATATTTCT 253

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/544,332
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/991,867
FILING DATE: 07-DEC-1992
APPLICANT: THALLEY, BRUCE S.
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/107,755
FILING DATE: 19-AUG-1993
APPLICANT: PADHYE, NISHA V.
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO 92/14818
FILING DATE: 12-FEB-1992
APPLICANT: FIRCA, JOSEPH R.
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,685
FILING DATE: 30-JAN-1992
APPLICANT: STAFFORD, DOUGLAS C.
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/657,584
FILING DATE: 19-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: Bencen, Gerard H.
REGISTRATION NUMBER: 35,746
REFERENCE/DOCKET NUMBER: UF114.C4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 1395 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-08-544-332-25

Query Match 4.8%; Score 48.8; DB 4; Length 1395;

Best Local Similarity 47.7%; Pred. No. 0.039;

Matches 207; Conservative 0; Mismatches 222; Indels 5; Gaps 2;

QY 475 TCTTTGCCGGGATGATATACACACCGATTATTTGGAGATTTTGATATTTCTTATGAT 534
DB 550 TTATATACAGACTTATAACACAAATAAATAATATGTTTATGATCTGATATGAC 491
QY 535 TTAACACTCCACCTAAGATGACCCCTATGGCGAATATATGATGATCTACAGCT 594
DB 490 TATAAAGCCCA--TAAACACTCAATAGTACCTCCCTAATCTAATTAATGCTTATTTATAC 433
QY 595 ATTTAAGACACCAAGAAATGATTTGGTATTAACCTAATATGAAATGAAATGATATT 654
DB 432 ATTTGATGCTACTCAGATATTTAAATGATATATTTAAATGAAATGAAATGATTTT 373
QY 655 TGTGCTATGGGTATTAATTTGACTGGTTTATATTCAGAAAATTTTCAAAGTGTGTTAGTC 714
DB 372 AATACCTTTTGAATTTACAATATGATGATGATGATGATGATGATGATGATGATGAT 313
QY 715 AAGATGATAAAGAAATTCACCTAATGATCTCATGTTAGTGATTTATATTTTAAATCAA 774
DB 312 TATATATAGTTTATTTCAACATTAATCTAATTTGATTTTAAATATATATTTATCT 253
QY 775 ATATTTGAAATTTCCGTACACCAATTTAAGTATTTGAGATGATTTTCTGAT 834
DB 252 ATGAGTTTATGGTGAATTTATGTCCT--AGTAGTTATAGAGGTGATATTTATTTCT 196
QY 835 GAATATAATAAGAAACCTTCGATTTTAAATAATTCAAATTTACAAAATATCTAGAAA 894
DB 195 AGATATGATATTAAGTAGTATTTGACGAATTTTAAATAATACATTTAGAAATTTTAA 136
QY 895 GATTGGGATATTAT 908

Db 135 AATAAAAGTTTAT 122
RESULT 15
US-08-480-604A-9
Sequence 9, Application US/08480604A
Patent No. 5736139
GENERAL INFORMATION:
APPLICANT: KINK, JOHN A.
APPLICANT: THALLEY, BRUCE S.
APPLICANT: PADHYE, NISHA V.
APPLICANT: FIRCA, JOSEPH R.
APPLICANT: STAFFORD, DOUGLAS C.
TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND
PREVENTION OF C. DIFFICILE DISEASE
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL, LLP
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,604A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/422,711
FILING DATE: 14-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/405,496
FILING DATE: 16-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/329,154
FILING DATE: 25-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/161,907
FILING DATE: 02-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/985,321
FILING DATE: 04-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/429,791
FILING DATE: 31-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: INGOLIA, DIANE E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: OPHD-01763
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 7101 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..7098
US-08-480-604A-9

Query Match 4.7%; Score 48; DB 2; Length 7101;
Best Local similarity 43.8%; Pred. No. 0.07;

| Matches | 310; Conservative | 0; Mismatches | 390; Indels | 8; Gaps | 2; |
|---------|-------------------|--|-------------|---------|----|
| QY | 195 | AATTGAATATTTAATGATCTTAAAAATTTATGATGATGTTATATTAGTCACCAAAATGTA | 254 | | |
| Db | 6105 | ATTTAAATATTTTGGCTATCATTAATGAAGATTTAGGAATGAAGAAGTGAAGAATCTC | 6164 | | |
| QY | 255 | TGGTTATGATTTGAGTCAATGATTGAAATFACAAAATATTGTAAACGAACAACGATT | 314 | | |
| Db | 6165 | ATATTTCTGGTATATTAATTTTCAATTAATAAATTTTACTATTTTGATGATTCATTTACAGC | 6224 | | |
| QY | 315 | TATTTATGGTATTAATGGTAAATCTTTGTTAGTAAATCAATATACACTTGTGTAATGAAATTGA | 374 | | |
| Db | 6225 | TGTAGTTGGATGGAAAGATTTAGACGATGGTTCAAAGTATATTTTGTGAAGAATACAGC | 6284 | | |
| QY | 375 | AGAAAAAGATATCAAAATTTATGTTTAAAAATCAATGAGTTCAGGACHTTGAATTTATTCATTC | 434 | | |
| Db | 6285 | AGNAGCATATAGGTTTCTCATTAATAAATGATGGTCAATATATTTTAAATGATGATGG | 6344 | | |
| QY | 435 | ACAAGGGATTAATCATCGTGATATAAACCCAGTAATATTTTCTTTCGCCGGGATGATAT | 494 | | |
| Db | 6345 | AAATATGCAAGTTGGATTTTGCCTATPAATATGATAAAGCTTCTTACTTCTGACTCTGG | 6404 | | |
| QY | 495 | AACAGAACCGGATATTGGAGATTTTGATATTGTTATGATTTAAACHTGCCACCTCAAGA | 554 | | |
| Db | 6405 | AAATATAGAATCTGGAGTACAAACCATAGATGACAAATATTTCTATATAGATGATATGG | 6464 | | |
| QY | 555 | TGAACCCCTATGGCGAAATATATTGATGATCTACAGGTATTTATAAGCACCACGAATT | 614 | | |
| Db | 6465 | TATAGTTCAAATTTGGTGTATTTGATPACTTCAGATGGATATAAATATTTTGCACCTGCTAA | 6524 | | |
| QY | 615 | GATCTCTGGTATAACTAAATATGAATATGAAATTGATATTGGTCATTGGGTATAAATTT | 674 | | |
| Db | 6525 | TACTGTAATGATAATATTTCGGACACAGCAGTTGAATATAGTGGTTTAGTTAGATGG | 6584 | | |
| QY | 675 | GACTGGTTTATATTCAGAAAATTTTCAAGTGTTTTAGTCAAGATGATAAAGAATTGAC | 734 | | |
| Db | 6585 | GGAGATGATATTT-----ATTTTGGNAGAACATATCAATTTGAGACTGGATGATATA | 6638 | | |
| QY | 735 | TAATGATCTCATGTTAGTGATTTATATTTTATTAATCAAAATATTTGAAATTTCCGGTAC | 794 | | |
| Db | 6639 | TGATATGGAATTTGAAGATGATAAATATTTTTCATTCACAGAACTAAAAGAATGCTCAA | 6698 | | |
| QY | 795 | ACCCA--ATTTAACTGATTTTGAAGATGAATATTTTGTGATGAATATAATAATGAAAC | 852 | | |
| Db | 6699 | AGGTATTAAATTTAATTTGATGATATAAATATATTTTGTATGAGAAGGCCATATAGAAC | 6758 | | |
| QY | 853 | TTGCATTTTAAAAAATTCAAATTTTCAAAAATATCCTAGAAAAGATGG | 900 | | |
| Db | 6759 | GGGTCTTATATCATTTTGAATAATAATTAATTTACTTTTATGAGATGG | 6806 | | |

Search completed: November 6, 1999, 02:12:41
Job time: 17958 sec